

République Algérienne Démocratique et Populaire  
Université Abou Bakr Belkaid– Tlemcen  
Faculté des Sciences  
Département d'Informatique

Mémoire de fin d'études  
pour l'obtention du diplôme de Master en Informatique

## *Thème*

# Segmentation des masses dans les images mammographiques

**Réalisé par :**

- Mr. Hamza CHEKKAF
- Mr. Ibrahim TOUIL

Présenté le 25 septembre 2012 devant la commission d'examination composée de MM.

- *Mohemed Amine CHIKH*..... (Encadreur)
- *Mohamed SERIARI*..... (Co-encadreur)
- *BENAZOUZ*..... (Examineur)
- *BENZIANE*..... (Examineur)
- *BENMOUNA* ..... (Examineur)

Année universitaire : 2011-2012

# Remerciements

Nous remercions **Dieu** de nous avoir accordé des connaissances de la science et de nous avoir aidé à réaliser ce travail.

Au terme de ce modeste travail nous tenons à remercier chaleureusement et respectivement tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste projet de fin d'étude.

Nos vifs remerciements vont tous d'abord à notre encadreur **M<sup>r</sup>. CHIKH MOHAMMED AMINE**.

Nous voulons également exprimer notre profonde gratitude pour notre Co-Encadreur **M<sup>r</sup>. SERIARI MOHAMMED** qui nous a encadrés tout le long de ce projet et qui nous a apporté beaucoup de connaissances dans ce vaste domaine des « traitements des images médicales ».

Tout notre respect et nos remerciements vont vers les membres du jury qui vont pleinement consacrer leur temps et leur attention afin d'évaluer notre travail, qui espérons le sera à la hauteur de leur attente.

Enfin, nos remerciements les plus sincères sont adressés à tous les professeurs, l'administration et le personnel de l'informatique qui ont contribué à forger nos connaissances et à assister notre formation.

# Dédicace

Louage à dieu, Seigneur des mondes ;

C'est toi que nous adorons et de toi que nous implorons secours ;

Que la paix et la bénédiction soit sur son dernier envoyé ;

A mes chers parents qui ont été présents, à chaque instant, leur  
irremplaçable et inconditionnel soutien m'a permis d'écartier les doute,  
de soigner les blessures et partager mes joies ;

A ma petite famille, mes frères et sœurs ;

A tout mes tantes et mes oncles ;

A mes cousins ;

A tout la famille CHEKKAF ;

A tout mes amis sans exception ;

**CHEKKAF HAMZA**

## DEDICACE

*A mes parents, aux être qui sont les plus chères au monde et auxquels je ne saurais jamais exprimer ma gratitude et ma reconnaissance en quelques lignes, Je les dédié ce modeste travail, que dieu le tout puissant les protège.*

*Pour ton amour, ton affection et ton soutien, pour ton courage et ton sacrifice, je te dédié, pour la deuxième et mille fois, ma très chère mère, un résultat modeste de la bienveillance et tes longues années de patience.*

*A mon très cher frère Abderrahmane et ma petite sœur Amina.*

*A mes grands parents, mes oncles, tantes et cousins.*

*Tous mes amis.*

*Et à tous ceux qui j'aime et qu'ils m'aiment... où qu'ils soient.*

*Ibrahim*

# Table des matières

	Page
<i>Chapitre I : Contexte Médicale</i>	
<b>I. Introduction</b>	6
<b>II. Anatomie du sein</b>	7
<b>III. Cancer du sein</b>	8
<b>IV. L'imagerie mammaire</b>	9
IV.1. Définition de l'image	9
IV.2. La mammographie	9
IV.2.1. De film radiologique aux clichés Numérisés	9
a- La numérisation indirecte	9
b- La mammographie numérique	9
IV.2.3. L'analyse des clichés mammographiques	11
IV.2.3.1. Densité mammaire	11
IV.2.3.2. Radio anatomie d'un sein normale	12
<b>V. Les opacités du sein</b>	12
V.1. Définition	12
V.2. L'analyse des opacités mammaire	13
V.3. Les propriétés des opacités dans les images mammographiques	15
<b>Conclusion</b>	17
<i>Chapitre II : L'analyse de texture</i>	
<b>I. Introduction</b>	19
<b>II. Définition de la texture</b>	20
<b>III. Types de textures</b>	20
<b>IV. L'analyse de texture</b>	21

IV.1. Méthode statistique	21
IV.1.1. Attributs statistique du premier ordre	22
IV.1.2. Attributs statistique du second ordre	23
<b>V. Conclusion</b>	<b>26</b>
<i>Chapitre III : Segmentation des images</i>	
<b>I. Introduction</b>	<b>28</b>
<b>II. Traitement d'image</b>	<b>28</b>
II.1. Images médicales	28
II.2. Définition de Traitement	28
<b>III. Segmentation des images</b>	<b>30</b>
III.1. Définition de la segmentation	30
III.2. Les différentes approches de segmentation	31
III.2.1. Segmentation basée sur les contours	32
a. Définition (un contour)	32
b. Méthode des contours actifs (snacke)	32
c. Méthode d'ensembles de niveau zéro (LEVEL SETS)	33
III.2.2. Approches de segmentation par région	35
III.2.2.1. Définition (Segmentation par région)	35
a. Méthodes par croissance de régions	35
b. Segmentation morphologique : LPE	36
c. K-means	37
III.2.3. Segmentation basée sur les pixels	38
III.2.4. Autres méthodes de segmentation	40
<b>IV. Conclusion</b>	<b>41</b>
<i>Chapitre IV : Implémentation</i>	
<b>I. Introduction</b>	<b>43</b>
<b>II. Synoptique de la démarche préconisée</b>	<b>43</b>
II.1. La base d'image	43
II.2. Prétraitement	44

<b>II.3. Calcule de La matrice de cooccurrence</b>	<b>44</b>
<b>II.4. Extraction des descripteurs de haralick</b>	<b>46</b>
<b>II.5. K-MEANS</b>	<b>47</b>
<b>III. Implémentation</b>	<b>48</b>
<b>III.1. Environnement matériel et logiciel de programmation</b>	<b>48</b>
<b>III.2. Présentation de quelques vues</b>	<b>48</b>
<b>III.3. Résultats et discussions</b>	<b>58</b>
<b>IV. Conclusion</b>	<b>59</b>
<b>Conclusion générale</b>	<b>60</b>
<b>Annexe</b>	
<b>Bibliographies</b>	

## Liste des figures

	Page
Figure 1 : Anatomie interne du sein	7
Figure 2 : Mammographie numérique(MN) et les clichés associés	10
Figure 3 : Différents types de densité mammaire	11
Figure 4 : Radio anatomie d'un sein normale sur une incidence face/oblique	12
Figure 5 : Exemple d'une masse (bénigne/maligne) dans la mammographie	13
Figure 6 : représentation schématique des formes de masses	14
Figure 7 : représentation schématique des contours de masses	14
Figure 8 : Exemples de différents types de masses	16
Figure 9 : Exemples de textures issues de l'album de Brodatz	20
Figure 10 : calcul de la matrice de co-occurrence d'une image	27
Figure 11 : Etapes du traitement d'images	29
Figure 12 : principe de contour actif	33
Figure 13 : principe de Level Sets	34
Figure 14 : Le processus d'un algorithme de croissance de région	35
Figure 15 : L'application de l'algorithme croissance de région sur les images mammographiques	36
Figure 16 : Principe de la Ligne de Partage des Eaux	37
Figure 17 : L'application de l'algorithme K-means	38
Figure 18 : Seuillage global pour $S=207$	39
Figure 19: Exemple de détermination du seuil	40
Figure 20 : Les principales étapes de l'algorithme de segmentation des masses du sein	43
Figure 21 : Fenêtre principal de l'application	49
Figure 22 : Information sur le logiciel	50
Figure 23 : fenêtre de traitement du logiciel	51
Figure 24 : fenêtre de traitement du logiciel	52
Figure 24 : fenêtre de paramètre « Contraste » de Matrice de cooccurrence (fenêtre de $3*3$ ).	52



Figure 25 : fenêtre de l'image texturée, paramètre « Variance » de Matrice de cooccurrence (fenêtre de 3*3).	53
Figure 26 : fenêtre de l'image texturée, paramètre «Cluster shade » de Matrice de cooccurrence (fenêtre de 3*3).	53
Figure 27 : fenêtre de l'image texturée, paramètre « Corrélation » de Matrice de cooccurrence (fenêtre de 3*3).	54
Figure 28 : fenêtre de l'image texturée, paramètre « Energie » de Matrice de cooccurrence (fenêtre de 3*3).	54
Figure 29 : fenêtre de l'image texturée, paramètre « Cluster Prominence » de Matrice de cooccurrence (fenêtre de 3*3).	55
Figure 30 : fenêtre de l'image texturée, paramètre « Homogénéité » de Matrice de cooccurrence (fenêtre de 3*3).	55
Figure 31 : segmentation sans calcul de descripteurs	55
Figure 32 : L'affichage de segmentation par K-means (fenêtre de 3*3).	57

## Liste des tableaux

	Page
Tableau 1 : Critères de distinction entre tumeurs bénignes/ malignes.	8
Tableau 2 : La classification BI-RADS de la densité mammaire	11
Tableau 3 : Tableau récapitulatif des opacités bénignes / malignes	15
Tableau 4 : présentation des paramètres de texture et leur signification.	25

## Liste d'abréviations

**μcas** : Micro calcifications

**MA** : Mammographie analogique

**MN** : Mammographie numérique

**CAD** : Systèmes d'aide au diagnostic

**1D** : Mono-dimensionnel

**3D** : Tridimensionnel

**NG** : Niveau de gris

**GLCM** : Gray Level Co-occurrence Matrix.

**ENER** : Energie

**CON** : Contraste (inertie)

**COR** : Corrélation

**VAR** : Variance

**HOM** : Homogénéité

**CSH** : Cluster Shade

**R/L** : Right/Left

**MLO** : Medio Latérale Oblique

**Fph** : Filtre passe haut

**Fpb** : Filtre passe bas

**Fi** : filtre identité

**LPE** : Ligne de Partages des Eaux

**EM** : Estimation-Maximisation

**textel** : texture élément

## Résumé

Dans un système d'aide au diagnostic (CAD) en mammographie numérique, la détection de masse est une tâche importante et délicate pour le diagnostic précoce du cancer du sein, cette difficulté se résume, pour une large part, dans la complexité des images mammographiques (densité mammaire) et la diversité des masses à segmenter côté contour et forme. Tenant compte de cette difficulté nous avons proposé un algorithme basé sur l'analyse statistique de la texture par les matrices de co-occurrences, les descripteurs de harralick et l'algorithme de clustering K-means.

**Mots Clés :** masse, Segmenté, Mammographiques, Analyse statistique, Texture, Matrice de co-occurrence, K-Means.

# Introduction Générale

## **Introduction Générale**

Le cancer du sein, problème de santé publique majeur, continue à être la première cause de mortalité chez les femmes à travers le monde. Il est pratiquement inexistant chez la fillette et très peu courant chez l'adolescente, bien qu'exceptionnel chez l'homme, souvent ignoré, il touche fréquemment les femmes de plus de 40 ans. L'incidence de cette lourde pathologie reste croissante en Algérie et il n'existe pas, jusqu'à l'heure actuelle, des programmes de dépistage organisés.

Différentes études ont confirmé que c'est la détection en stade précoce des cancers qui peut en améliorer le pronostic. Dans ce cas la mammographie reste la technique de référence incontournable pour l'exploration du sein, la plus performante en matière de surveillance et de dépistage précoce du cancer du sein. Elle permet de mettre en évidence des anomalies telles que les opacités et les calcifications qui peuvent traduire des lésions malignes. Cependant tous les radiologues reconnaissent la difficulté de l'interprétation des mammographies qui s'accroît encore par le type de tissus du sein examiné.

Dans le même contexte se situe notre travail qui portera sur la segmentation des masses dans les images mammographiques dans l'objectif d'extraire les zones suspectes pour faciliter le diagnostic de ces images ultérieurement : La question n'est pas de se substituer au médecin, mais de lui proposer des outils performants qui l'aident dans son travail d'analyse des clichés de mammographie.

Dans ce mémoire nous allons exposer les différentes étapes de notre processus proposé pour la segmentation des masses dans les images mammographiques.

- **Plan du mémoire :**

Le mémoire est organisé comme suit :

### ***Chapitre 1 : Contexte Médicale***

Dans ce chapitre nous décrivons tout d'abord l'architecture de la glande mammaire, et nous avons défini le cancer de sein et comparé les deux grandes catégories de tumeurs. Mettre en relief la technique d'imagerie mammaire par excellence, la mammographie radiologique. Une analyse fine et approfondie porte sur le type d'anomalie mammaire « les opacités ».

## ***Chapitre 2 : L'analyse de la Texture***

Ce chapitre est consacré à La texture, qui est presque omniprésente dans les images, joue donc un rôle important en analyse d'images non seulement dans les applications de segmentation mais aussi de classification et de caractérisation.

## ***Chapitre 3 : La segmentation***

Nous décrivons les différentes approches et quelques méthodes de segmentation, et nous nous intéressons par une méthode basée région qui est l'algorithme de clustering « k-Means ».

## ***Chapitre 4 : Implémentation***

Ce chapitre présente le cœur de ce mémoire : il est dédié à la détection de type de lésions mammaire « les opacités », signes d'alerte d'un cancer du sein. Leur détections est indispensable pour un système d'analyse automatique en mammographies. Et nous exposons un processus de segmentation des masses dans les images mammographiques.

Enfin, nous concluons notre travail en résumant les principales contributions que nous avons apportées dans le domaine de la segmentation des images mammographiques.

# Chapitre I : Le contexte Médicale



## **I. Introduction**

Dans ce chapitre, nous avons présenter l'approche médicale de l'imagerie du sein, pour situer le contexte dans lequel l'application que nous cherchons à développer va être utilisée. Nous présentons l'anatomie du sein, le cancer du sein et des changements bénignes et malignes qui peuvent se produire, mais cette approche n'est pas exhaustive et est donnée uniquement dans le but de faciliter la compréhension de la problématique pour en déduire des méthodes de détection. Nous décrivons ensuite les généralités sur l'imagerie mammaire et les lésions radiologiques qui le constituent : nous nous intéressons par les masses. Nous terminons ce chapitre par présenter en détail les masses du sein dans la mammographie numérique.

## II. Anatomie du sein

Organe pair et symétrique de forme hémisphérique, situé en avant du thorax, entre la troisième et la cinquième côte, au-dessus du muscle grand pectoral. Une coupe longitudinale passant par le mamelon (Figure 1). La glande mammaire se divise en 15 à 20 unités glandulaires indépendantes, lobes mammaires, divisés en lobules par des cloisons conjonctives, qui s'achèvent en douzaines de minuscules bulbes, acini, base de tout raisonnement physiopathologique du sein. Les lobes, lobules et bulbes sont reliés entre eux par les canaux galactophores, qui se terminent par un canal unique s'abouchant au niveau du mamelon, situé au centre d'une zone pigmentée, l'aréole. La plus grande partie du sein est constituée de tissu adipeux qui comble l'espace situé entre les différentes structures du sein. En période d'allaitement, les lobules fabriquent du lait et les canaux le transportent vers le mamelon où il est tété par le bébé. Les tissus mammaires baignent dans des liquides lymphatiques, qui permettent d'éliminer les déchets, les cellules mortes et autres débris. La circulation lymphatique du sein est importante dans la mesure où elle va déterminer les lieux où l'on retrouvera des métastases lors de cancer.[1]

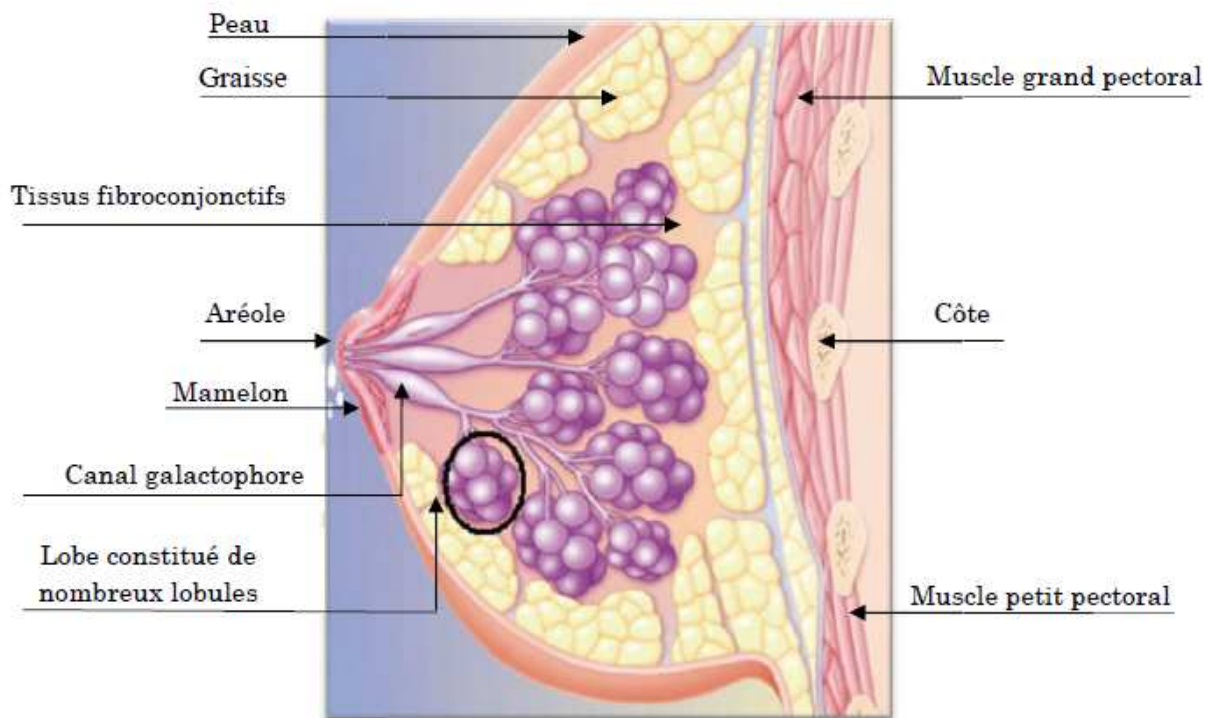


Figure 1 : Anatomie interne du sein

### III. Cancer du sein

Le cancer du sein est une tumeur maligne se développant à partir des cellules constituant la glande mammaire. Les cellules malignes se multiplient de manière anarchique et forment la tumeur. Lorsqu'un cancer du sein n'est pas traité, les cellules tumorales se propagent localement et envahissent les organes de voisinage (extension locale puis extension régionale). Elles peuvent également se propager par voie sanguine ou lymphatique (extension générale) pour atteindre les organes situés à distance (métastases).

Les organes les plus fréquemment atteints par les métastases sont les os, les poumons, le foie et le cerveau.[1]

Il y a deux grandes catégories de tumeurs sont connues :

- Les tumeurs bénignes
- Les tumeurs malignes ou cancers.

<b>Tumeurs bénignes</b>	<b>Tumeurs malignes</b>
Bien limitée	Mal limitée
Encapsulée	Non encapsulée
Histologiquement semblable au tissu d'origine	Plus ou moins semblable au tissu d'origine (dédifférenciation, différenciation aberrante)
Cellules régulières	Cellules irrégulières (cellules cancéreuses)
Croissance lente	Croissance rapide
Refoulement sans destruction des tissus voisins	Envahissement des tissus voisins
Pas de récurrence locale après exérèse complète	Récurrence possible après exérèse supposée totale
Pas de métastase	Métastase(s)

**Tableau 1** : Critères de distinction entre tumeurs bénignes/ malignes.

## IV. L'imagerie mammaire

### IV.1. Définition de l'image :

L'image est une représentation d'un objet par la peinture, la sculpture, le dessin, la photographie, le film, etc. C'est aussi un ensemble structuré d'informations qui, après affichage sur l'écran, ont une signification pour l'œil humain.

- **Représentation des images numériques :**

une image numérique est une matrice de pixels repérés par leur coordonnées (x,y). S'il s'agit d'une image couleur, un pixel est codé par 3 composantes (r,g,b) (chacune comprise au sens large entre 0 et 255), représentant respectivement les "doses" de rouge, vert et bleu qui caractérisent la couleur du pixel. S'il s'agit d'une image en niveau de gris, il est codé par 1 composante comprise au sens large entre 0 et 255, représentant la luminosité du pixel.

### IV.2. La mammographie (imagerie par rayons X)

La mammographie est la technique la plus performante en matière de surveillance et de dépistage précoce du cancer du sein. La qualité de l'image mammographie est caractérisée par trois paramètres physiques : la résolution, le contraste et le bruit.

Selon Bermond, près de 40% des cancer « ratés » à la mammographie sont dus à une mauvaise qualité des clichés. Il s'agit le plus souvent d'un défaut de positionnement du sein[2].

#### IV.2.1. De film radiologique aux clichés Numérisés

Deux méthodes de reproduction de l'image mammographie à partir du film radiologique :

**a- La numérisation indirecte :** s'effectue à partir d'un film et une camera CCD ou un numériseur de film à balayage laser (scanner optique) [3]. Cependant, une radiographie scannée ne contiendra jamais plus d'information que l'épreuve originale. En plus elle conserve tous les défauts de l'image de départ (bruit, dynamique, artefacts de développement,...).

**b- La mammographie numérique :**

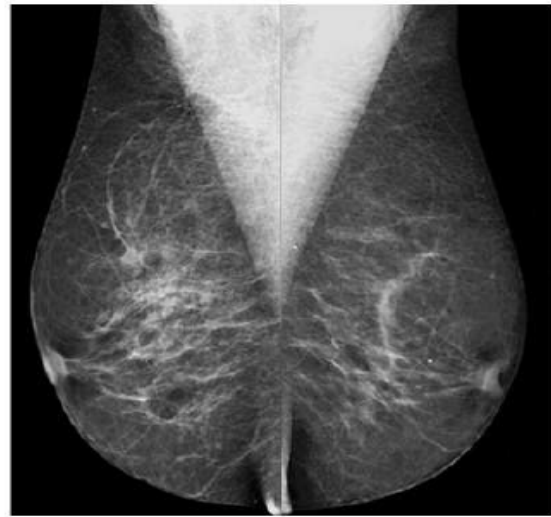
Le couple écran film est remplacé par un détecteur qui convertit proportionnellement une intensité de rayons X transmise par le sein en un signal électronique. La gamme dynamique d'un détecteur numérique est beaucoup plus grande que celle du couple écran film, ce que

justifie la détection de certains détails de l'ordre de microns. Cette technologie présente de multiples avantages :

- Manipulation des images sur des écrans d'ordinateur de haute résolution en fonction des besoins (zoom, inversion contraste, fenêtrage,..) avec une dose d'irradiation moins importante.
- Cette technologie numérique accélère considérablement l'acquisition des images et élimine le temps consacré au développement de la pellicule, ce qui permet d'accroître l'offre d'examens des patients.[1]



(a)-Mammographie numérique



(b)- Clichés associés

Figure 2 : Mammographie numérique(MN) et les clichés associés

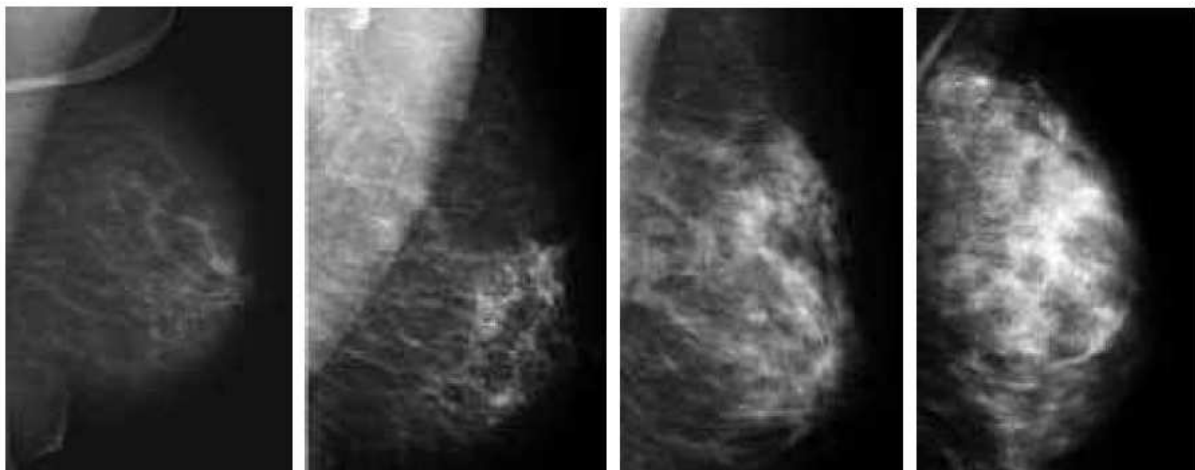
### IV.2.3. L'analyse des clichés mammographiques

#### IV.2.3.1. Densité mammaire

A la ménopause, les seins sont souvent « déshabités » et donc tout à fait radio transparents. Les images tumorales s'y inscrivent alors avec une très particulière netteté. La classification BI-RADS de l'*American College of Radiology* permet de classer les seins en quatre groupes pour lesquelles le risque de rater un cancer en raison du caractère hétérogène du tissu fibroglandulaire augmente lorsque l'on passe de la catégorie 1 à la catégorie 4 (Figure 3).[1]

Type de la densité mammaire	Type
Le sein est presque entièrement graisseux : seins clairs.	1
Il y a des opacités fibroglandulaires dispersées.	2
Le tissu mammaire est dense et hétérogène.	3
Le tissu mammaire est extrêmement dense.	4

Tableau 2 : La classification BI-RADS de la densité mammaire



a) Type 1

b) type 2

c) type 3

b) type 4

Figure 3 : Différents types de densité mammaire

### IV.2.3.2. Radio anatomie d'un sein normale

Soulignons qu'il est extrêmement délicat de définir une normalité des clichés mammographiques. L'aspect de la glande mammaire est en effet extrêmement variable selon la patiente, son âge, et par rapport au cycle menstruel.

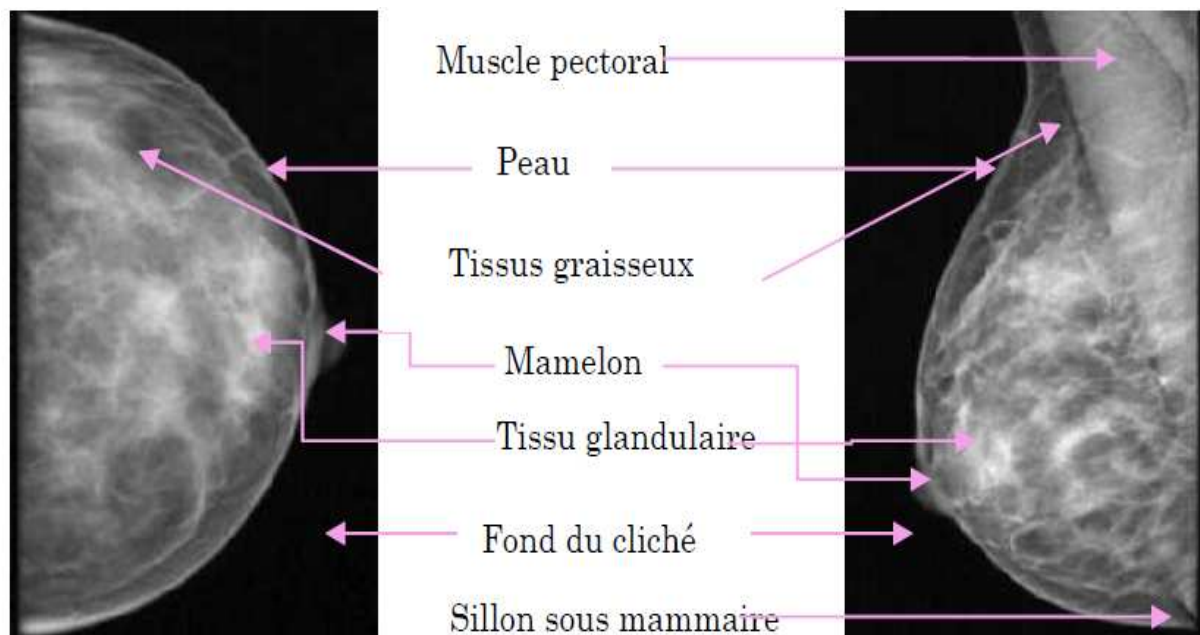


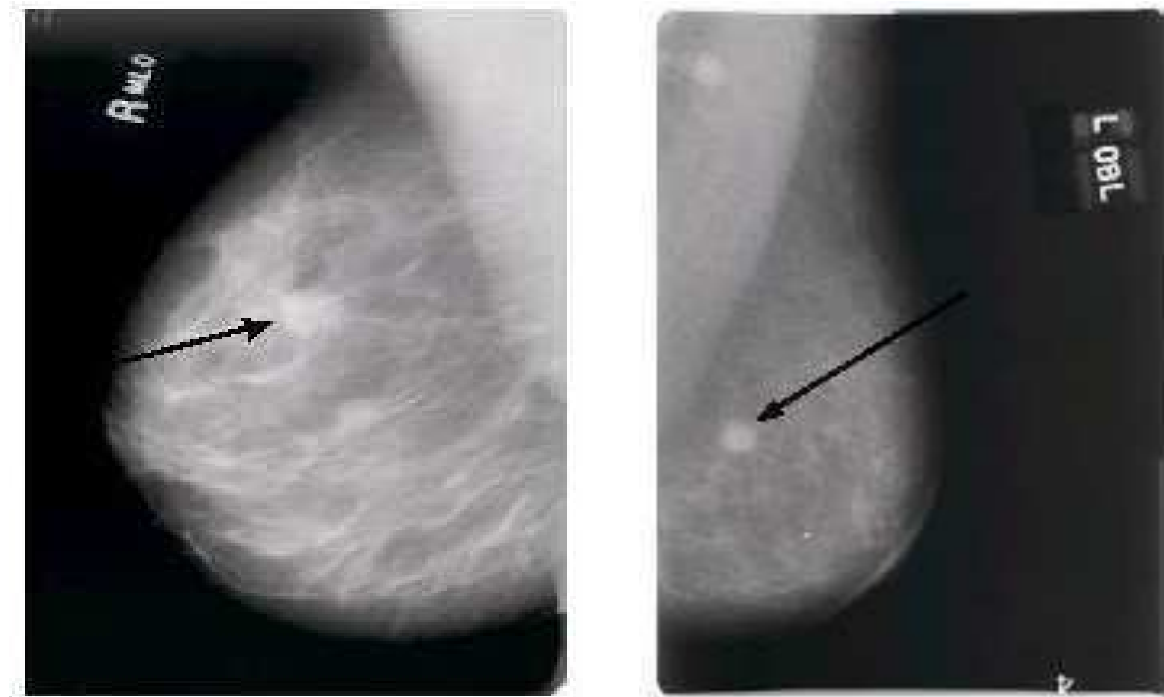
Figure 4 : Radio anatomie d'un sein normale sur une incidence face/oblique

## V. Les opacités du sein

### V.1. Définition

Une opacité correspond à une plage de surdensité anormale (sur l'exemple de la figure 5, nous avons indiqué l'opacité par une flèche), car les surdensités normales sont nombreuses sur un cliché mammographique. Une sur-densité anormale ne se distingue pas d'une sur-densité normale par un critère précis, mais par une combinaison de différentes caractéristiques : taille, densité, contour, forme, texture... C'est l'expérience qui permet au radiologue de distinguer une opacité sur une mammographie. Une surdensité sur plusieurs clichés effectués sous plusieurs incidences implique une forte présomption en faveur de l'opacité.

Une opacité traduit une anomalie du tissu conjonctif ou épithélial. Elle sera donc aisément visible au niveau d'une zone grasseuse et beaucoup plus difficilement perceptible dans une zone dense de tissu conjonctif. [4]



(a) Mammographie montrant une masse bénigne. (b) Mammographie montrant une masse maligne.

Figure 5 : Exemple d'une masse (bénigne/maligne) dans la mammographie

## V.2. L'analyse des opacités mammaire

Il existe quatre critères radiologiques :

- **La taille** : elle est variable de quelque millimètre à plusieurs centimètre. Cependant la taille ne prédit pas le caractère malin, sauf sur des clichés successifs lorsque l'on voit la taille augmenter régulièrement.
- **La densité** : ces masses sont en général dense (plus blanches) que le tissu mammaire adjacent.



- **La forme** : on peut envisager cinq formes de masse :

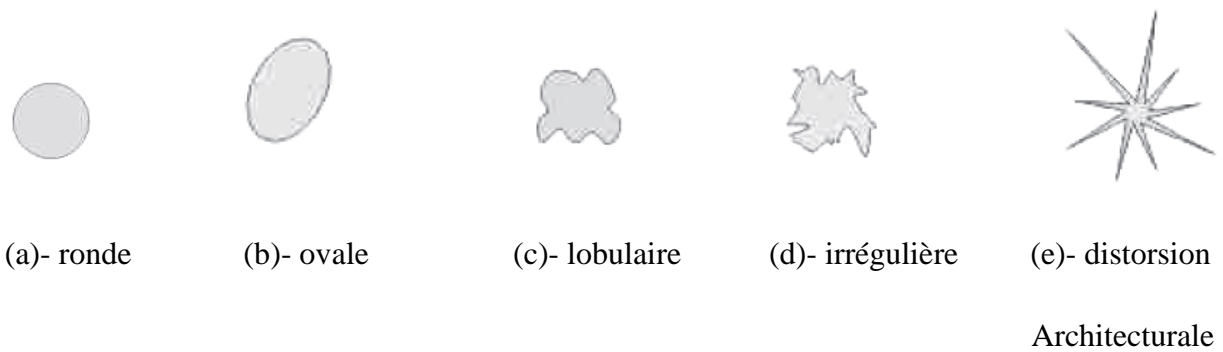


Figure 6 : représentation schématique des formes de masses

- **Les contours** : c'est le meilleur critère d'études des opacités. La figure 7 montre cinq types de contours masses :

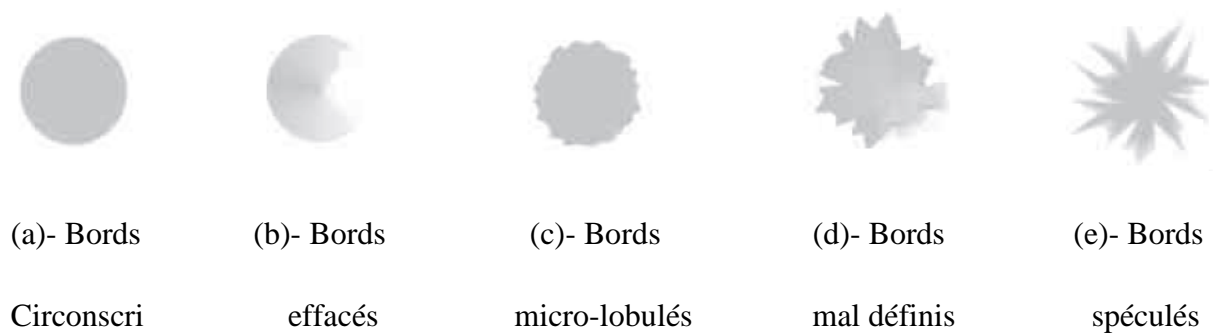


Figure 7 : représentation schématique des contours de masses

En générale, une opacité régulière, ronde, ovale ou lobulée avec contour nets n'a aucun caractère inquiétant car elles sont a priori des lésions bénignes. Cependant, cette règle n'est pas absolue, certain cancer peut prendre les mêmes caractères. Cependant une opacité anormale à centre hyperdense ; des limites peu nettes et irrégulier sont des signe de suspicion de malignité.[1]

Le tableau ci-dessous (Tableau 3) résume des opacités (bénigne ou maligne) ;

en fonction des critères radiologiques étudiés précédemment :

<b>Paramètres radiologiques</b>	<b>Critères de classification d'opacité</b>	
	<b>Opacité négatives</b>	<b>Opacités «suspecte »positives</b>
<b>Taille</b>	En dessous ou nettement en dessous de 10mm	10mm ou plus
<b>Morphologie</b>	Ronde, elliptique, sans lobulation ni rectitude localisée	Polylobée, avec des rayons de courbure variée, une rectitude localisée
<b>Contour</b>	Parfaitement net	Zone de flou limité, microlobulations

Tableau 3 : Tableau récapitulatif des opacités bénignes / malignes

### V.3. Les propriétés des opacités dans les images mammographiques

Rappelons qu'une opacité est une anomalie focale de densité, traduisant une lésion des tissus conjonctifs ou épithéliaux. Sur une image mammographique, elle apparaît comme une région de surdensité légèrement plus lumineuse qu'un tissu normal.

Au niveau d'une zone grasseuse elle apparaît sous le meilleur contraste, donc aisément visible, tandis qu'au niveau d'une zone dense de tissu conjonctif, elle sera plus difficilement perceptible (Figure 8).

Leur nature (bénignes ou malignes) est définie par une combinaison de différentes caractéristiques : taille, densité, contour, forme, ...ext.

Une masse circonscrite, habituellement ronde, lobulaire ou ovale avec des bords nets est souvent bénigne. Tandis qu'une masse de forme irrégulière ayant des frontières mal définies / flous ou spéculés (lignes radiaires prenant naissance sur le contour de la masse) est suspecte et représente un risque élevée de malignité. Les exemples de la figure 8 illustrent les différents cas possible d'une opacité mammaire.

Un œil non expert différencie très difficilement les surdensités anormales présentes sur les clichés, d'ailleurs c'est l'expérience qui permet au radiologue de distinguer une opacité d'une sur-densité normale. [5]

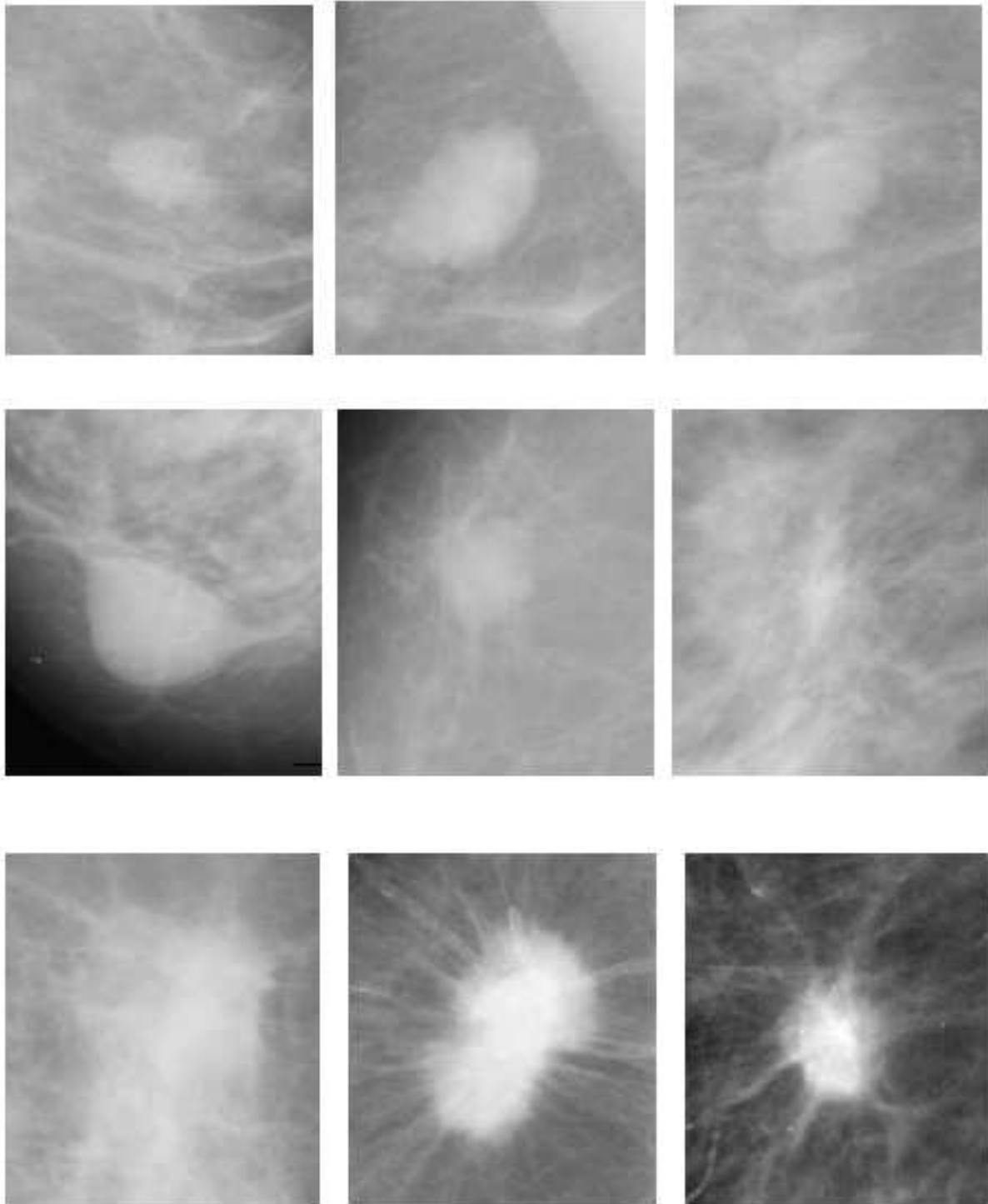


Figure 8 : Exemples de différents types de masses

## **VI. Conclusion**

Dans ce chapitre, nous avons introduit quelques notions générales concernant l'anatomie du sein, l'apport de la mammographie, les différentes modalités d'imagerie mammaire. En raison du faible contraste de beaucoup de lésions cancéreuses et de la similitude flagrante de quelques lésions avec le tissu du sein, l'interprétation du cliché de mammographie devient une tâche très difficile qui demande une très grande habileté et une forte concentration.

# Chapitre II : L'analyse de Texture

## I. Introduction

La notion de texture en traitement d'image est d'autant plus délicate à aborder que les définitions que l'on en donne sont multiples. Ces définitions sont souvent liées à un aspect particulier mais sont rarement génériques et la quantité importante d'approches utilisées pour l'analyse des textures témoigne de l'absence d'une définition précise. L'objet de ce chapitre préliminaire est tout d'abord de faire une synthèse des différentes interprétations de la notion de texture.

Bien que la notion de texture soit naturelle pour l'être humain, elle résiste depuis longtemps à toute tentative de définition. On peut s'en approcher en disant qu'une texture est une zone de l'image qui présente certaines caractéristiques d'homogénéité qui la font apparaître comme une zone unique. Nous pouvons aussi la décrire comme étant un ensemble de primitives de taille et de forme variables, présentant une organisation spatiale particulière.

[06]

## II. Définition de la texture

Une texture peut être fine ou grosse. Ou bien, elle est retrouvée par la distribution des couleurs. Elle est peut-être détectée par sa primitive - le textel (texture élément), dans ce cas, la texture est une répétition des textels [07].

on n'a pas pu trouver une définition formelle de ce qu'est la texture et c'est la raison pour l'abondance des méthodes pour déterminer ou différencier des textures. On se contente donc de trouver un modèle adéquat pour l'étude à mener.

En conséquence, les définitions de texture proposées varient selon les domaines de recherche et les conceptions des auteurs.

## III. Types de textures

Il existe deux types de textures: les textures périodiques définies par la répétition d'un motif de base et les textures dites aléatoires qui ont un aspect anarchiques mais qui apparaissent homogènes vis-à-vis de leur environnement. Brodatz a établi un catalogue de texture à des fins de synthèse d'images texturées [08]. Ce catalogue constitue aujourd'hui une référence en matière de base de test pour les chercheurs travaillant sur la texture. Il s'agit de 112 textures (sable, herbe, bulles, eau, bois, etc.) ayant chacune seize variantes différentes. Ces images sont classiquement utilisées pour valider les différentes méthodes d'analyse de texture.

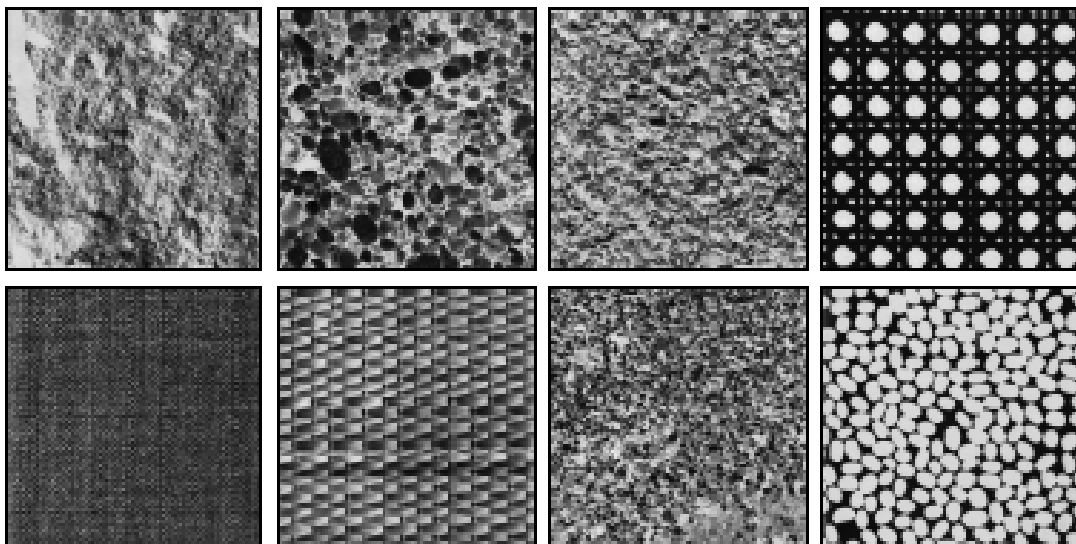


Figure 9 : Exemples de textures issues de l'album de Brodatz

## IV. l'analyse de texture

L'analyse de texture regroupe un ensemble de techniques permettant de quantifier les différents niveaux de gris présents dans une image en termes d'intensité et de distribution dans le but de calculer un certain nombre de paramètres caractéristiques de la texture à étudier.

Les applications possibles de l'analyse de texture sont la segmentation, la classification ou la caractérisation de textures homogènes. Comme il existe deux grandes classes de textures, l'analyse peut se faire selon deux classes de méthodes: les méthodes structurelles et les méthodes statistiques [06]. Les méthodes structurelles permettent de décrire la texture en définissant les primitives et les "règles" d'arrangement qui les relie. Elles sont donc applicables sur les textures aléatoires. Alors que les méthodes statistiques étudient les relations entre un pixel et ses voisins et définissent des paramètres discriminants de la texture en se basant sur des outils statistiques. Généralement, elles sont utilisées pour caractériser des structures fines, sans régularité apparente. Plus l'ordre de la statistique est élevé et plus le nombre de pixels mis en jeu est important.

### IV.1. Méthode statistique

Ces méthodes se basent aussi sur les relations entre les pixels et leurs voisins en termes de variation de couleur, les attributs extraits de ces méthodes permettent de caractériser tout type de texture, même les textures fines et sans régularité apparente.

De nombreux descripteurs statistiques sont utilisés dans la discipline de traitement d'images texturées couleur. On trouve entre autres [07].

- Les statistiques d'image.

- Les histogrammes d'image.

- Les matrices de co-occurrences chromatiques.

- Les histogrammes des sommes et des différences.

- Les matrices de longueurs de plages.

- Etc.

Ces descripteurs peuvent être divisés en plusieurs catégories selon leur ordre. L'ordre du descripteur dépend du type d'interaction spatiale entre les pixels considérés. Par exemple, pour les catégories des descripteurs statistiques du premier ordre. Les matrices de co-



occurrences son quant-à-elles des descripteurs d'ordre 2 car leur calcul considère des couples de pixels.

#### **IV.1.1. Attributs statistique du premier ordre**

##### **➤ Statistiques d'image**

De nombreuses statistiques peuvent directement être extraites d'une image afin de caractériser la texture couleur qu'elle contient. Les statistiques les plus courantes sont [07] :

La moyenne des composantes couleurs,

La variance ou l'écart-type qui constitue une estimation du degré de dispersion des composantes couleurs autour de leur moyenne,

La dissymétrie qui estime le degré d'asymétrie des valeurs des composantes couleurs par rapport à leur moyenne,

L'aplatissement qui estime le degré de concavité ou convexité des composantes couleurs par rapport à leur moyenne.

Cependant, ces statistiques n'exploitent que l'information de couleur des pixels et négligent les relations spatiales entre ces derniers. La caractérisation des textures couleurs à partir de ces attributs est donc peu performante.

##### **➤ Histogrammes d'image**

Les histogrammes font également partie des descripteurs statistiques du premier ordre. Ils décrivent la distribution de niveaux des composantes couleurs de l'image. Cette description peut être effectuée de différentes façons [07] :

L'histogramme peut être implanté selon une structure tridimensionnelle (3D) composée de cellules caractérisant chaque couleur. Si les composantes couleurs sont codées sur 8 bits, l'histogramme de l'image contiendra alors  $256^3$  cellules. Chacune de ces cellules contient en fait le nombre de pixels de l'image caractérisés par la couleur considérée.

Une autre description consiste quant à elle à considérer indépendamment les trois composantes couleurs par le biais des histogrammes mono-dimensionnels (1D) des images-composantes.

Les attributs qui se sont souvent extraits de ces histogrammes pour discriminer les différentes classes sont :

- La médiane, qui partitionne l'histogramme en deux parties d'égale population,
- Le mode, qui correspond au maximum de l'histogramme,
- L'intervalle interquartile, qui est une caractéristique de dispersion autour de la médiane.

Tout comme pour les statistiques d'image, l'inconvénient majeur des histogrammes est qu'ils n'exploitent que l'information de couleur et négligent les relations spatiales entre les pixels. Il est donc nécessaire de considérer des attributs statistiques d'ordre supérieur pour une analyse de texture plus efficace.

#### IV.1.2. Attributs statistique du second ordre

##### ➤ Matrice de co-occurrence

La matrice de co-occurrence est largement utilisée dans l'analyse de texture. Elle est très facile à mettre en œuvre et donne de bons résultats sur plusieurs types de texture. Dans la plupart des applications, les images utilisées sont représentée en niveaux de gris et la matrice de co-occurrence est connue sous le nom GLCM (Gray Level Co-occurrence Matrix).

La matrice de co-occurrence  $C$  d'un vecteur de déplacement  $(\Delta x, \Delta y)$  d'une image  $I$  de la taille  $(m*n)$  est définie par la formule :

$$C_{\Delta x, \Delta y}(x, y) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{si } I(p, q)=i \text{ et } I(p+\Delta x, q+\Delta y)=j \\ 0, & \text{sin on} \end{cases}$$

Pour améliorer la performance de la méthode, nous pouvons réduire la taille de la matrice. Pour le faire, nous pouvons grouper des niveaux gris de l'image par les techniques de partitionnement de données (clustering).

La Figure 10 montre la façon dont est calculée la matrice de co-occurrence du vecteur de déplacement  $(1,0)$ .

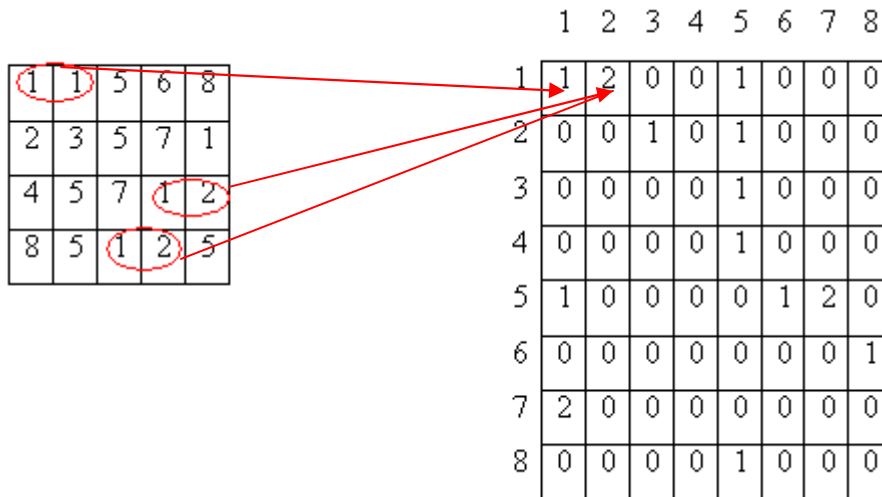


Figure 10 : calcul de la matrice de co-occurrence d'une image

Le choix du vecteur de déplacement est toujours très important pour la réussite de la méthode. Normalement, nous voulons obtenir des matrices de co-occurrence de plusieurs vecteurs de déplacement sur des directions et des distances différentes.

Le problème est la combinaison des indices des matrices pour que nous puissions utiliser cette méthode avec des orientations et des échelles différentes. À côté de ce problème, nous voulons réduire le nombre de matrices de co-occurrence calculées. Pour faire cela, nous devons évaluer l'importance des vecteurs de déplacement par rapport aux types de textures. Plusieurs études ont été menées pour déterminer une distance ou une orientation optimale. En pratique, une distance courte donne généralement de bons résultats [09],[10].

La masse d'informations sur cette matrice est trop grande et nous ne pouvons pas retirer directement des remarques utiles pour l'analyse de texture. Quatorze indices intermédiaires (Table 4) sont proposés par Haralick en 1973 [11]. Ces indices réduisent l'information contenue dans la matrice de co-occurrence et permettent une meilleure discrimination entre les différents types de textures.

En outre, nous pouvons réduire le nombre d'indices utilisés dans les quatorze caractéristiques de Haralick. Pour l'indépendance de la taille d'image et pour l'expression en termes de probabilité, nous utilisons la matrice de co-occurrence normalisée pour calculer

Paramètre de Texture	Signification	Formule Mathématique
Second moment angulaire (énergie)	faible lorsque les $p(i,j)$ ont des valeurs très proches et une grande valeur lorsque certaines valeurs sont grandes et d'autres petites.	$ENER = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} p(i, j)^2$
Contraste (inertie)	Ce paramètre a une valeur numérique importante si les $p(i,j)$ sont concentrés hors diagonale.	$CON = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i - j)^2 * p(i, j)$
Corrélation	Ce paramètre a une grande valeur quand les valeurs sont uniformément distribuées dans la matrice de co-occurrence et une faible valeur dans le cas contraire.	$COR = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)(j - \mu_y) p(i, j)}{\sigma_x \sigma_y}$
Variance	Ce paramètre mesure la distribution des niveaux de gris autour de la moyenne.	$VAR = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i - MOY)^2 * p(i, j)$
Homogénéité locale	Ce paramètre indique une mesure de l'uniformité des niveaux de gris de l'image.	$HOM = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} \frac{p(i, j)}{1 + (i - j)^2}$
Cluster Shade	Comme la variance, le groupe de nuance mesure la distribution des niveaux de gris autour de la moyenne, avec une capacité de discrimination élevée à l'ordre 3.	$CSH = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i + j - MOY)^3 * p(i, j)$

Tableau4 : présentation des paramètres de texture et leur signification.

*Presentation of some textural parameters and their significance.*

- Il existe aussi d'autres descripteurs et pour voir plus suivi l'annexe A.

## V. Conclusion

Le domaine de l'étude de la texture est encore loin d'être bien exploré. En effet toutes les méthodes de synthèse de texture considèrent la texture comme un vecteur de paramètres statistiques. Il ressort de la présentation sus faite que de tels vecteurs ne suffisent pas pour rendre compte de manière précise de la texture d'une image.

Nous réaffirmons qu'il faudrait tenir compte de l'origine physique de la texture, des paramètres statistiques, de l'approche structurale, parler le même langage que dans la mesure où la définition de la texture devient moins approximative et où dans l'analyse de l'image l'aspect objectif l'emporte sur le subjectif. Il nous semble que l'avenir de la maîtrise de l'image réside dans la compréhension plus approfondie de sa texture, donc dans sa définition plus complète et cohérente.

# **Chapitre III :**

# **Segmentation des**

# **images**

## **I. Introduction**

La segmentation d'images est une étape incontournable dans tout processus d'analyse d'image. C'est un traitement de bas niveau qui précède l'étape de mesure, de compréhension et de décision. Son objectif consiste à partitionner l'image en régions connexes et homogènes au sens d'un critère d'homogénéité difficile à définir surtout dans le cas de régions texturées.

Nous aborderons dans ce chapitre les principales méthodes de segmentation, accompagnées de brèves descriptions et quelques illustrations.

## **II. Traitement d'image:**

### **II.1. Images médicales:**

L'imagerie médicale représente l'ensemble des techniques permettant d'obtenir des images à partir de différents types de rayonnements (ultrasons, rayons X, etc.).

Au cours des deux dernières décennies le traitement des images médicales est devenu un outil fondamental dans la médecine, et le traitement des images cérébrales en est un paradigme.

Le développement rapide des techniques d'acquisition des images médicales a permis au personnel médical d'avoir une grande variété de données appartenant à des images de différentes modalités, à titre d'exemple l'échographie ultrasonore, le scanner, la tomographie par émission de positons, la tomographie par émission de photon unique et l'imagerie par résonance magnétique IRM.

### **II.2. Définition de Traitement :**

C'est un ensemble des opérations relatives à la collecte, à l'enregistrement, à l'élaboration, à la modification, à l'édition, . . . de données.

Mettons de cotés les termes enregistrement et édition. Le principe général du traitement d'image est donc à quelques détails près toujours le même un système reçoit des images, y applique un traitement, et produit une information de nature liée à l'application visée.

[12]

Une source de rayonnement envoie des ondes sur un objet, qui sont ensuite réfléchies et collectées par un capteur. Le capteur transforme ces ondes en un ensemble de points. Ces points sont traités et une information est produite en sortie du système. On peut résumer le traitement d'image en quatre étapes principales :

➤ Acquisition des images

Mise en œuvre des processus physiques de formation des images suivis d'une mise en forme pour que ces images puissent être traitées par des systèmes informatiques.

➤ Traitement des images

Son but : améliorer ces images lorsqu'elles possèdent du bruit ou des défauts.

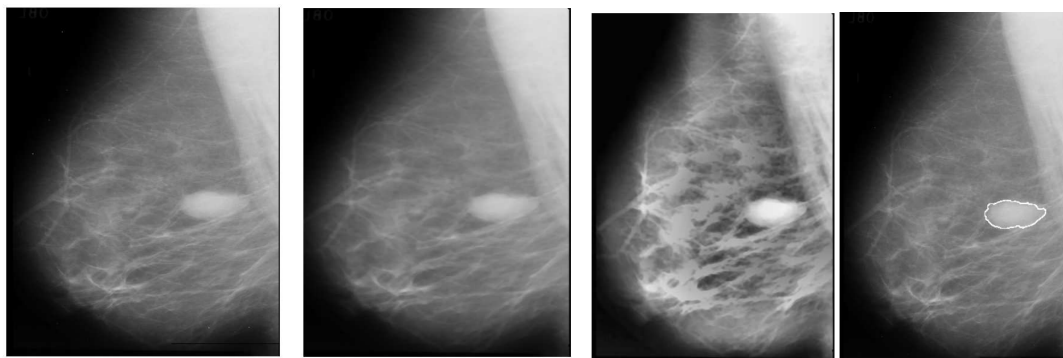
➤ Segmentation des images

Son but : construire une image symbolique en générant des régions homogènes selon un critère défini à priori.

➤ Analyse des images

Consiste à extraire des paramètres ou des fonctions représentatives de l'image ou des régions.

Nous pouvons illustrer ces étapes par un exemple simple.[01] (figure 11)



**Acquisition**

**Filtrage**

**Segmentation**

Figure 11 : Etapes du traitement d'images



### III. Segmentation des images :

#### III.1. Définition de la segmentation

La segmentation d'image est un traitement de bas-niveau et une des étapes critiques de l'analyse d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis. Chaque groupe de pixels forme alors une région. Une région est donc un ensemble connexe de pixels ayant des propriétés communes (intensité, texture,...) qui les différencient des pixels des régions voisines. Si le nombre de régions est égal à deux, la segmentation est appelée aussi binarisation. [13].

De façon plus précise, on peut définir la segmentation comme étant une partition d'une image  $I$  en  $n$  ensembles  $R_i$  appelés régions tels que :

- 1)  $\bigcup_{i=1}^n C_i = I$
- 2)  $\forall i, j \in \{1, \dots, n\}^2$  et  $i \neq j \Rightarrow C_i \cap C_j = \emptyset$
- 3)  $\forall i \in \{1, \dots, n\}^2$   $C_i$  est connexe
- 4)  $\forall i, j \in \{1, \dots, n\}^2$   $P(C_i) = \text{Vrai}$ ,
- 5)  $\forall i, j \in \{1, \dots, n\}^2$   $C_i$  est adjacent à  $C_j$  et  $i \neq j \Rightarrow P(C_i \cup C_j) = \text{faux}$

Où le prédicat  $P$  est utilisé pour tester l'homogénéité.

La première condition signifie que l'image  $I$  est partitionnée en  $n$  classes. Et la deuxième explique que tous les classes sont disjointes deux à deux.

La troisième étape, la quatrième et la cinquième imposent à chaque pixel d'une classe de satisfaire à la même propriété au sens du prédicat  $P$ . Le prédicat  $P$  n'est plus vrai pour la réunion de deux classes adjacentes.

Donc, la segmentation permettant une analyse de données en regroupant les pixels formants l'image à classifier de telle sorte les pixels appartenant à une classe soient plus similaires entre eux que ceux des classes différents, on affecter à chaque entité de la scène traité une étiquette indiquant son appartenance à une classe particulière. L'entité utilisée caractérise généralement un pixel ou un ensemble de pixel, tandis que l'étiquette constitue un thème choisi par l'utilisateur.

Il n'y a pas de méthode unique de segmentation d'une image, le choix d'une technique est lié :

- à la nature de l'image :
  - éclairage non homogène, reflets,
  - présence de bruit, de zones texturées,
  - contours flous, en partie occultés,
- aux opérations situées en aval de la segmentation :
  - localisation, mesure, calcul 3D,
  - reconnaissance des formes, interprétation,
  - diagnostic, contrôle qualité,
- aux primitives à extraire :
  - contours, segments de droite, angles,...
  - régions, formes,
  - textures,
- aux contraintes d'exploitation :
  - complexité algorithmique, fonctionnement en temps réel,
  - taille de la mémoire disponible en machine.

### **III.2. Les différentes approches de segmentation**

Essentiellement, l'analyse d'images a pour but l'extraction de l'information caractéristique contenue dans une image. Cette information peut prendre la forme, la couleur, le contour, la texture,... Donc, il est nécessaire de procéder tout d'abord à la segmentation de la lésion. Cette segmentation est généralement abordée selon deux modèles complémentaires : l'approche région et l'approche frontière.

L'approche frontière regroupe les techniques associées à une variation d'intensité : détection de contours. Cependant, les contours obtenus sont rarement connexes, donc elle doit faire face au problème *de fermeture de contours*.

L'approche région considère des groupements de pixels ayant des propriétés communes, chaque pixel est affecté à une région unique après une partition de l'image.

Toutefois, il n'existe pas d'algorithme spécifique pour déterminer les régions et les frontières dans une image. D'ailleurs les chercheurs ont compromis que la segmentation idéale n'existe pas, la bonne technique de segmentation sera donc celle qui permettra d'arriver à une bonne interprétation [14].

Une panoplie de techniques est proposée dans la littérature où chacune présente des avantages et des inconvénients. Parmi les nombreuses approches de segmentation d'images qui existent, nous présentons quelques techniques les plus citées dans la littérature appliquées à la mammographie, à savoir les méthodes basées sur pixels, l'approche région ainsi que celles basées sur les contours.

### III.2.1. Segmentation basée sur les contours:

La détection de contour est souvent le premier problème qu'on rencontre en traitant une image. La difficulté augmente avec l'importance de bruit présent.

**a. Définition (un contour) :** un contour peut approximativement être défini comme une frontière entre deux régions où l'intensité des pixels change brusquement.

Généralement l'utilisation d'un tel opérateur de contour se combine avec un seuillage et comme étant ce dernier est généralement imparfait, on obtient, d'une part, des contours qui ne limitent pas les régions fermées. Donc on doit faire recours à *des algorithmes de fermetures des contours*. D'un autre part, les zones de fortes variations ne correspondent pas forcément à un contour d'objet. Alors un post-traitement est nécessaire pour analyser les différents contours obtenus.

#### **b. Méthode des contours actifs (snake)**

L'idée de cette méthode est de déplacer les points pour les rapprocher des zones de fort gradient tout en conservant des caractéristiques comme la courbure du contour ou la répartition des points sur le contour ou d'autres contraintes liées à la disposition des points.

Au démarrage de l'algorithme, le contour est disposé uniformément autour de l'objet à détecter puis il va se rétracter pour en épouser au mieux la forme.

A chaque itération, l'algorithme va tenter de trouver un meilleur positionnement pour le contour pour minimiser les dérives par rapport aux contraintes utilisées.

L'algorithme s'arrêtera lorsqu'il ne sera plus possible d'améliorer le positionnement ou simplement quand le nombre maximum d'itérations aura été atteint. On utilise les notions d'énergies interne et externe pour caractériser respectivement la forme du contour et son positionnement sur l'image en tenant compte des lignes de gradient.

Pour une étude plus détaillée on peut se référer au [15].

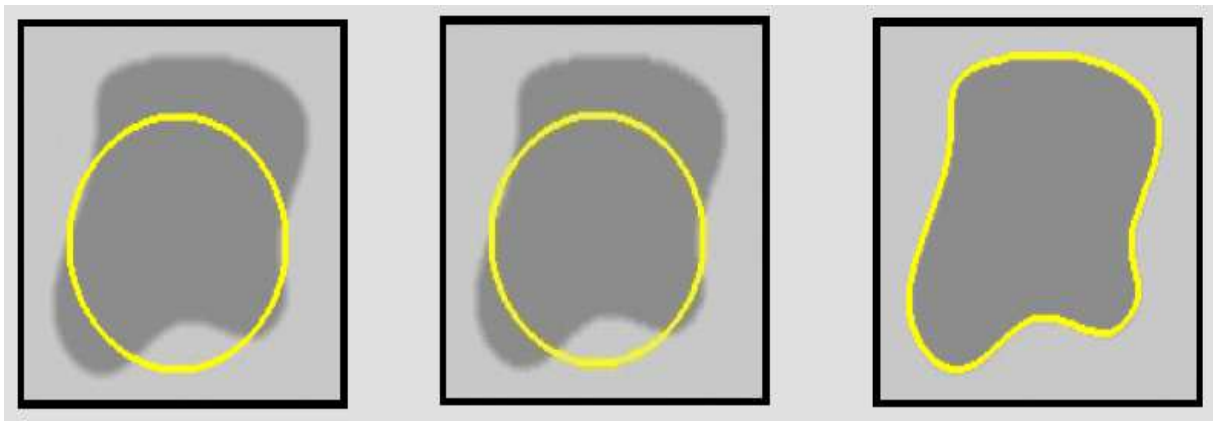


Figure 12 : principe de contour actif

### c. Méthode d'ensembles de niveau zéro (LEVEL SETS)

La méthode des Ensembles de Niveau Zéro est une méthode de simulation numérique utilisée pour l'évolution des courbes et des surfaces dans les domaines discrets [16].

L'idée de base de la méthode des level sets est de considérer une courbe (ou interface) en mouvement comme le niveau zéro d'une fonction de dimension plus élevée. Pour une courbe en 2D, cette interface ( $\Psi$ ) est l'intersection d'une hypersurface (de dimension 3) avec un plan.

Les points définissant cette interface vont se déplacer vers la normale à une vitesse  $F$  selon l'équation suivante [22]:

$$\Psi_{t+1} + F |\nabla \Psi_t| = 0$$

Cette vitesse  $F$  est composée de trois termes : un terme constant (similaire à la force

d'inflation utilisée dans les modèles déformables), un terme dépendant de la courbure locale en chaque point et un terme dépendant de l'image (dans notre cas, les fronts de l'image).

Le schéma numérique de l'équation de déplacement de l'interface est décrit par l'équation :

$$\Psi_{n+1} = \Psi_n - dt * k_1(x, y) * (U_n - \varepsilon K) * |\nabla \Psi_t| \varepsilon \in [0, 1]$$

Avec :

$U_n(m, \sigma) = \pm 1$ , fonction d'appartenance définissant la zone ou l'objet à rechercher.

$K = \nabla \cdot \frac{\nabla \Psi}{|\nabla \Psi|}$ , courbure locale en chaque point de l'interface.

$kI(x, y)$ , critère d'arrêt dépendant de l'image de gradient.

L'initialisation est réalisée avec une ou plusieurs formes de départ.

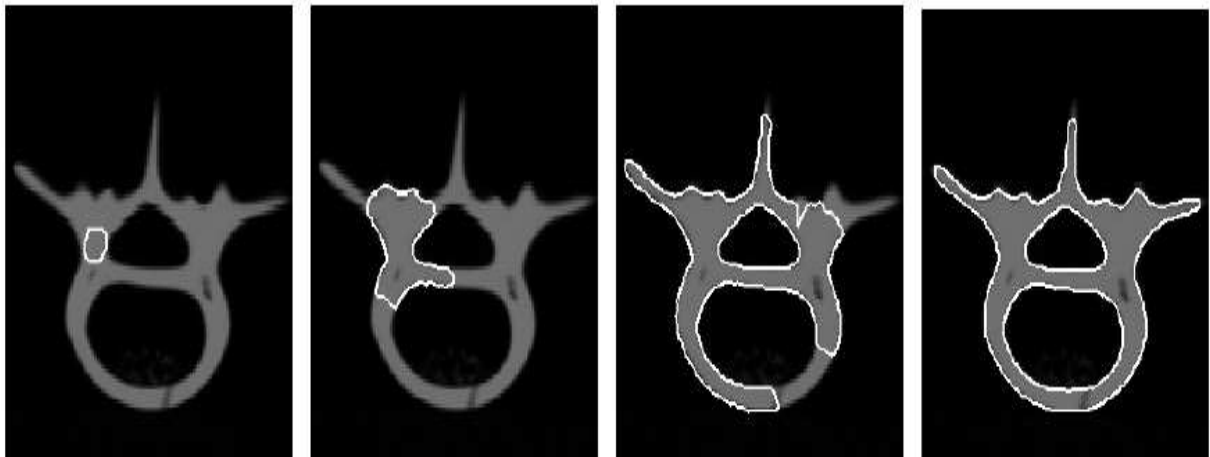


Figure 13 : principe de Level Sets

L'avantage principal de cette méthode est la possibilité de gérer automatiquement le changement de topologie de la courbe en évolution.

Et ses limitations :

- Problème de divergence : Méthode de résolutions.
- Interpolation : perte de matière.
- Cout de calcul : Réinitialisation de la fonction de distance.

### III.2.2. Approches de segmentation par région

L'approche par régions consiste à regrouper des points selon des propriétés communes.

**III.2.2.1. Définition (Segmentation par région) :** Globalement, elle peut être définie comme une partition d'une image  $I$  en une ou plusieurs régions  $R_1, \dots, R_n$  telles que :

$$I = \bigcup_{i=1}^n R_i \text{ et } R_i \cap R_j = \emptyset \quad \text{pour } i \neq j$$

#### a. Méthodes par croissance de régions

Le principe de ces méthodes est de réunir de façon itérative un ensemble de points connectés en une région de plus en plus large, en fonction de critères d'homogénéité. Pour définir une région, on définit un *germe* dans la région d'intérêt qui sert comme un point de départ pour l'agrégation (Figure 14).

a- Cependant un inconvénient à marquer à cette méthode c'est qu'elle est récursive: risque de débordements (pile).

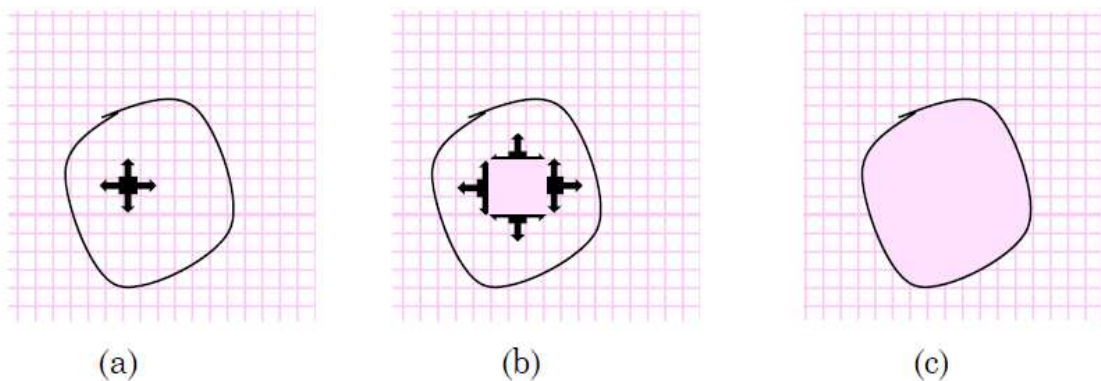


Figure 14 : Le processus d'un algorithme de croissance de région. (a) début de processus, le point noir présente le germe et les flèches la direction de croissance. (b) la croissance de pixels après quelques itérations. (c) le résultat de la segmentation.

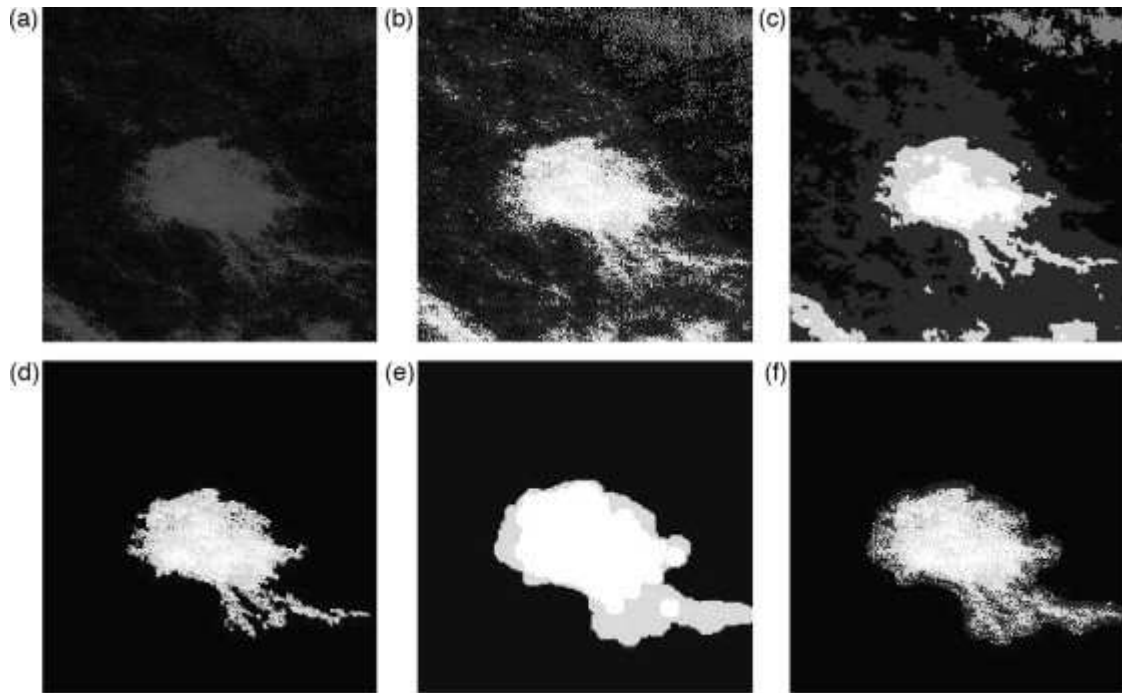


Figure 15 : L'application de l'algorithme croissance de région sur les images mammographiques [17]

### b. Segmentation morphologique : LPE

La Ligne de Partage des Eaux (LPE) est une technique de segmentation d'image assez répandue. Cet outil consiste à chercher les contours des bassins versants dans l'image à segmenter. Après avoir déterminé les contours de ces bassins qui forment les minima locaux, on obtient une carte de régions. L'inconvénient majeur de cette technique réside dans sa sensibilité au nombre de germes (généralement les minima locaux) qui s'avère, souvent, très important et engendre donc une sur-segmentation de l'image.

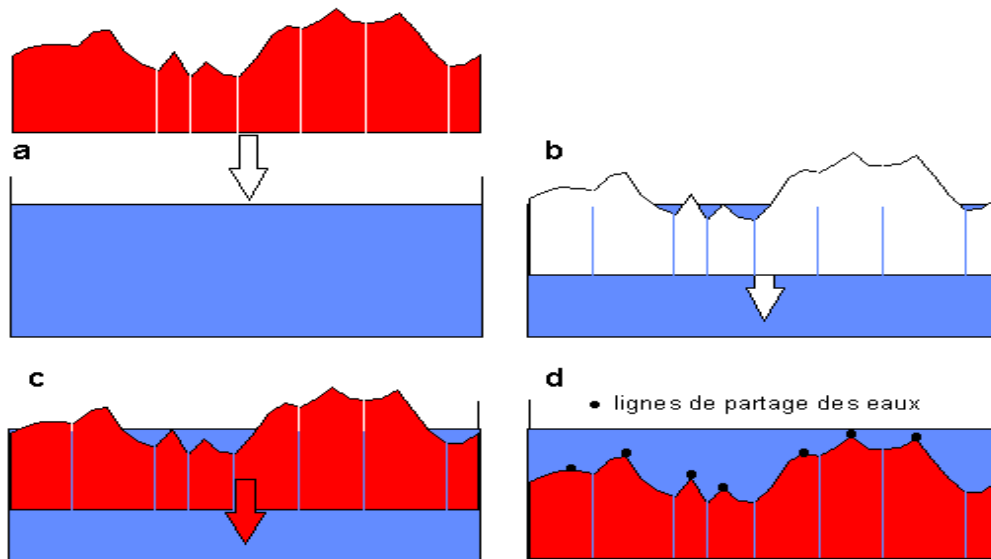


Figure 16 : Principe de la Ligne de Partage des Eaux

### c. K-means :

Cette méthode consiste à rassembler les pixels en  $K$  groupes (clusters).  $K$  étant un paramètre préfixé qui détermine le nombre de régions. Et après avoir déterminé les paramètres (couleur par exemple) de toute région, chaque point est affecté au centre le plus proche, pour recalculer ensuite le paramètre de chaque région, jusqu'à ce qu'elles soient stables.

#### ▪ Algorithme K-means:

L'algorithme des K-Means est l'une des techniques de clustering non supervisée les plus utilisées. La méthode consiste à placer aléatoirement dans l'espace  $K$  "centroïdes" afin de déterminer  $K$  clusters. On affecte à chacun de ces "centroïdes" les objets les plus proches, puis on calcule la position moyenne des objets associés aux "centroïdes" que l'on déplace en ce point. Les opérations d'affectation d'objet et de déplacement du centroïde sur la moyenne repérées jusqu'à ce que chaque centroïde ait atteint une position stable.

Malgré sa simplicité, cet algorithme se révèle efficace. Toutefois il est nécessaire de prédéterminer le nombre de catégories et la position de départ des prototypes qui a un impact sur le découpage en classes. Les principales étapes de cet algorithme sont:[18]

- Choix aléatoire de la position initiale des  $K$  clusters.
- (Ré-)Affecter les objets à un cluster suivant un critère de minimisation des distances (généralement selon une mesure de distance euclidienne).
- Une fois tous les objets placés, recalculer les  $K$  centroïdes.



- Répéter les étapes 2 et 3 jusqu'à ce que plus aucune ré-affectation ne soit faite.



Image original

Image segmentée avec  $K=5$ 

( Le nombre de classe est 5)

Figure 17 : L'application de l'algorithme K-means[19]

Le principal avantage de cette méthode est que le nombre de régions est connu au préalable (il n'y aura pas de problème de sur-segmentation ou sous-segmentation), mais l'inconvénient réside en la difficulté de déterminer le nombre de clusters et l'incohérence des régions (deux objets éloignés peuvent appartenir à la même région) [20].

### III.2.3. Segmentation basée sur les pixels

La segmentation basée sur le pixel travaille sur des histogrammes de l'image, on citera essentiellement la méthode du seuillage. La segmentation est inhérente puisque les régions sont déduites du seuillage.

#### a. **Seuillage:** (Thresholding)

Le seuillage est une méthode faisant partie de l'approche basée sur le pixel, même si d'autres références classent cette méthode dans l'approche région de la segmentation, du fait qu'elle met en évidence les régions d'une image.

Un seuil est une valeur numérique correspondant à un paramètre de l'image (exemple : le niveau de gris), et l'histogramme des niveaux de gris est utilisé afin d'en déduire le seuil.

Le seuillage peut être :

-**Global** : un seuil pour toute l'image

-**Local** : un seuil pour une portion de l'image

-**Adaptatif** : un seuil qui s'ajuste selon les images, ou parties de l'image [21]

Le seuillage le plus simple est le seuillage global, il a pour principe :

Soit  $p$  un pixel et  $V$  une fonction définissant le niveau de gris et  $S$  un seuil :

$$V(p) = \begin{cases} 0 & \text{si } V(p) < S \\ 1 & \text{si } V(p) \geq S \end{cases}$$

Le résultat sera évidemment une image binaire.

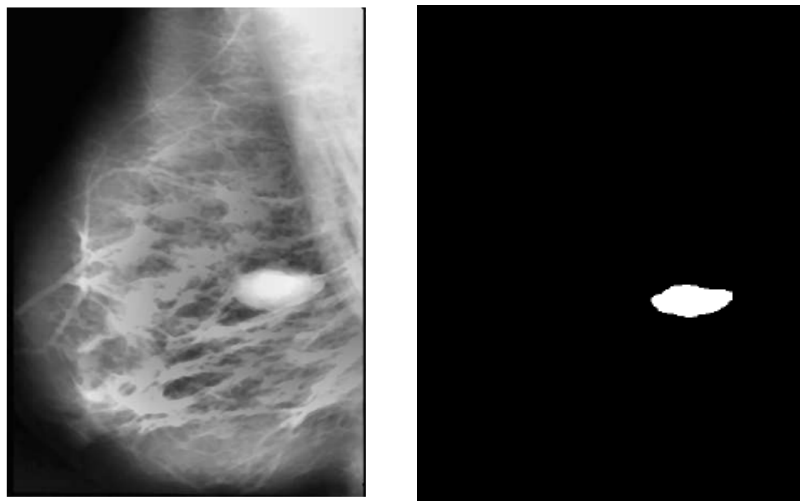


Figure 18 : Seuillage global pour  $S=207$

Le seuillage global est certes rapide, mais il ne donne pas de pixels contigus, et le bruit peut donner de faux éléments.

Le seuillage local détermine pour chaque pixel un seuil en fonction de la luminosité de son voisinage.

Le seuillage adaptatif se résume en la séparation de l'image en sous images et de traiter chacune avec son propre seuil (Le choix de la dimension des sous-images est critique). [21].

Dans tous les cas, le problème majeur reste la détermination du seuil, il existe quelques méthodes de détermination telles que :

- La recherche des vallées de l'histogramme, en considérant que les vallées significatives correspondent aux pixels de la frontière des régions.
- La segmentation de l'histogramme (développée par Otsu) qui ne s'applique que dans le cas de segmentation de l'image en deux régions.

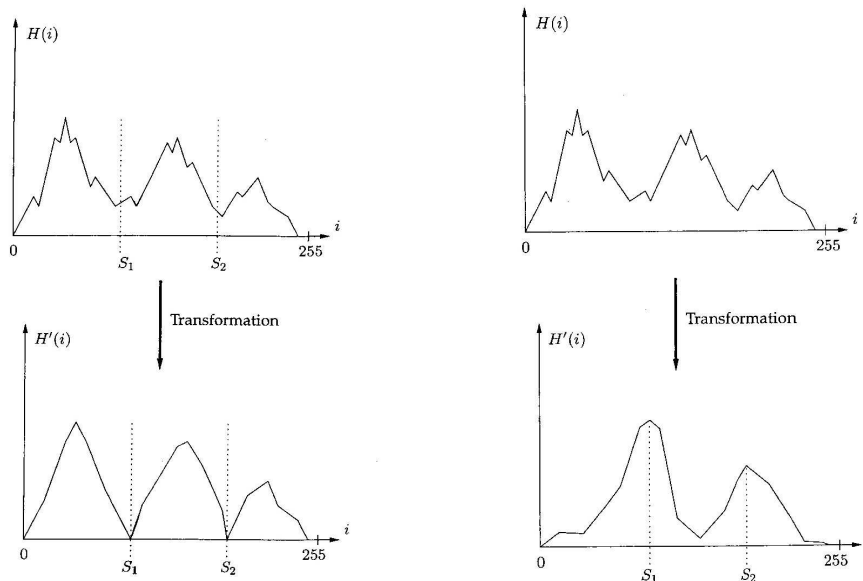


Figure 19 : Exemple de détermination du seuil.

### III.2.4. Autres méthodes de segmentation

D'autres méthodes de segmentation sont utilisées, à l'image de :

- Opérateurs morpho-mathématique (érosion, dilatation, ouverture, fermeture, squelettisation).
- Méthodes hybrides [18] :
  - Unification entre l'approche région et l'approche contour.
  - Fusion à l'aide de régions floues des régions précédemment segmentées.
  - Segmentation d'images basée sur les statistiques des rangs des niveaux de gris.
  - Système hybride de segmentation non supervisée d'images par des Chaînes de Markov cachées (CMC).
  - Segmentation d'images basée sur un modèle fonctionnel.

## IV. Conclusion

La segmentation est une étape primordiale en traitement d'image et du fait qu'il n'existe pas de méthode unique de segmentation, il est difficile de définir, de manière absolue, une « bonne » segmentation. La segmentation n'est pas une fin en soi, sa qualité est fonction des résultats obtenus par les traitements situés en aval qui utilisent les primitives extraites.

Ce chapitre dresse brièvement un état d'art des principales méthodes de segmentation. Chacune possède ses avantages et ses inconvénients : le problème majeur de la segmentation basée sur les contours est la fermeture des contours [22], d'autre part, la segmentation basée sur les régions peut s'avérer « gourmande » en calcul. En l'absence de contextes particuliers,

**Puisque la qualité du résultat est principalement Subjective et que la finalité d'utilisation de la segmentation obtenue influe sur la perception de qualité de ce résultat, le choix d'une technique plutôt qu'une autre est assez difficile.**

# Chapitre IV :

# Implémentation

## I. Introduction

L'objectif de ce travail est de proposer une méthode de segmentation des masses dans les images mammographies. Le but de cette méthode est de faciliter la détection et l'interprétation des masses du sein.

Ce chapitre dresse la synoptique de la démarche préconisée ainsi la présentation des différentes interfaces de l'application réalisée.

## II. Synoptique de la démarche préconisée :

La synoptique de notre démarche est décrite par la Fig.4.22 :

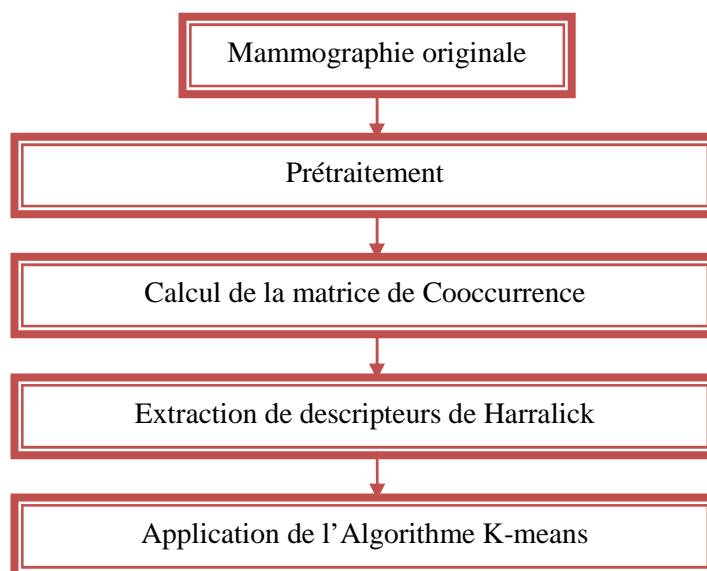


Figure 20 : Les principales étapes de l'algorithme de segmentation des masses du sein

### II.1. La base d'image:

Dans ce travail, nous utilisons la base d'image mini-Mias (Mammography Image Analysis Society) [23]. MIAS est une organisation des groupes de recherches sur les mammographies du royaume uni (United Kingdom) qui a développé une base de données de mammographies numériques avec une résolution spatiale de 1024\*1024 pixels pour chaque image, cette base contient 322 images sont réparties en 207 images normales, 38 images contiennent des masses et 169 contiennent d'autres anomalies. Dans cette étude nous avons sélectionné 23 images.

## II.2. Prétraitement :

Dans une chaîne de traitement d'images, la phase de segmentation constitue une étape clé. Lorsque les images sont d'origine naturelle, on est confronté à la présence de bruit dans ces images et à l'existence de contours peu marqués rendant difficile la formation des régions ou la localisation des frontières. Il faut dès lors procéder à un traitement destiné à réduire le bruit et renforcer les contrastes locaux. Du fait du caractère non stationnaire du processus d'acquisition d'image, la qualité de cette dernière est généralement de faible contraste et pour cela nous avons besoin des filtres pour résoudre ce problème. Nous proposons ici deux filtres, médian et Passe haut.

Pour la suppression des bruits, le filtre médian, appliqué sur un voisinage de  $3 \times 3$  permet d'éliminer les valeurs aberrantes de l'image.

L'intérêt du filtre passe haut est d'isoler les contours en augmentant les hautes fréquences et atténuant les basses. Ce filtre ne conservera, des images, que les zones de fortes variations de luminosité ou de chrominance. Nous avons appliqué ce filtre avec une taille de masque de  $5 \times 5$

## II.3. Calcul de La matrice de cooccurrence :

La méthode des matrices de cooccurrence demeure la méthode de référence pour l'analyse de la texture naturelle, elle est largement utilisée dans le monde du traitement d'image. Elle présente une grande performance et donne de bons résultats sur la plupart des types d'images ; ce qui justifie notre choix de la confronter avec une autre méthode d'analyse aussi performante «K-means».

La méthode des matrices de cooccurrence exploite les relations qui existent entre les pixels de l'image dans une direction  $\theta$  et une distance donnée  $d$ . Il s'agit en fait de la probabilité de trouver, si on part d'un pixel de valeur  $i$ , un pixel de valeur  $j$  quand on se déplace en suivant un vecteur dans la direction  $\theta$  d'une distance  $d$ . [24]

Un important problème de ces matrices est leur manipulation. En effet, elles sont de taille imposante et utilisent un grand espace de mémoire, puisque le nombre d'éléments qui les composent est de  $N_g * N_g$  ( $N_g$  étant le nombre de niveaux de gris de l'image). C'est pour cette raison même que l'on procède à une étape de normalisation. [25]

En plus ces matrices contiennent une masse très importante d'informations difficilement manipulable. C'est pour cela qu'elles ne sont pas utilisées directement mais à travers des

mesures dites *indices de texture*. Ces indices réduisent l'information contenue dans la matrice de cooccurrence et permettent une meilleure discrimination entre les différents types de textures. La question principale lors de la génération des paramètres texturaux concerne le choix de la taille de la fenêtre d'analyse et du vecteur de déplacement, ceux sont les paramètres sur lesquels repose la réussite de la méthode. La difficulté à surmonter lors de l'application de cette technique réside justement dans le choix de ces paramètres car ils varient en fonction du type d'images et de textures. Pour ce qui est du vecteur de déplacement, si toutes les directions sont étudiées avec différents pas de déplacement, un grand nombre de matrices de cooccurrence doit être calculé. Ceci conduit à une grande complexité calculatoire. Plusieurs études ont été menées pour déterminer une distance ou une orientation optimale en concluant, qu'en pratique, une distance courte donne généralement de bons résultats. [26].

La matrice de co-occurrence  $C$  d'un vecteur de déplacement  $(dx,dy)$ , d'une image  $P$  de taille  $M \times N$  est définie par la formule suivante:

1- Initialisation :

$$\sum_{x=1}^{N_g} \sum_{y=1}^{N_g} C(x, y) = 0$$

Où  $N_g$  est la valeur maximal du niveau de gris de l'image  $P$

2- Remplissons de la matrice :

$$\sum_{x=1}^{N_g} \sum_{y=1}^{N_g} \sum_{i=1}^N \sum_{j=1}^M \begin{cases} C(x,y)=1+C(x,y) & \text{si } p(i,j)=i \text{ et } p(i+dx,j+dy)=j \\ C(x,y)=C(x,y) & \text{sinon} \end{cases}$$

Généralement on ne calcule pas la matrice de co-occurrence sur l'image d'origine, mais plutôt sur l'image ayant subi une réduction des niveaux de gris. On réduit généralement le nombre de niveau de gris de 256 à 16,32 ou bien 64. Cette réduction du nombre de niveaux de gris permet de de garder l'information nécessaire relative aux régions pertinentes de l'image, et surtout de supprimer tous les problèmes de bruits et autres déformations éventuelles qui sont présentes sur l'image.



## II.4. Extraction des descripteurs de haralick:

Quatorze (14) descripteurs de textures basées sur le calcul des matrices de cooccurrences ont été proposés par haralick, Julesz, Gagalowicz et Weska [10], dans ce qui suit nous ne présentons que sept (07) indices que nous utilisons dans notre travail :

### ✚ Le Contraste

Mesure les variations locales des niveaux de gris. Si elles sont importantes

$$CON = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i - j)^2 * p(i, j)$$

### ✚ Le Clustershad

Le groupe de nuance mesure la distribution des niveaux de gris autour de la moyenne, avec une capacité de discrimination élevée à l'ordre 3.

$$CSH = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i + j - MOY)^3 * p(i, j)$$

### ✚ La Variance

Ce paramètre mesure la distribution des niveaux de gris autour de la moyenne.

$$VAR = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i - MOY)^2 * p(i, j)$$

### ✚ La Corrélation

Ce paramètre a une grande valeur quand les valeurs sont uniformément distribuées dans la matrice de co-occurrence et une faible valeur dans le cas contraire.

$$COR = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)(j - \mu_y) p(i, j)}{\sigma_x \sigma_y}$$

### ✚ L'Energie

faible lorsque les  $p(i,j)$  ont des valeurs très proches et une grande valeur lorsque certaines valeurs sont grandes et d'autres petites.

$$ENER = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} p(i, j)^2$$

### 🚩 Le Cluster Prominence

$$C P N = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i + j - MOY)^4 * p(i, j)$$

### 🚩 L'Homogénéité

Est d'autant plus élevé que l'on retrouve souvent le même couple de pixels, ce qui est le cas lorsque le niveau de gris est uniforme ou quand il y a une périodicité spatiale

$$HOM = \sum_{i=0}^{N_g} \sum_{j=0}^{N_g} \frac{p(i, j)}{1 + (i - j)^2}$$

## II.5. K-MEANS :

C'est l'un des algorithmes les plus connus en classification non supervisée. Il vise à produire un partitionnement des pixels issus de deux classes différentes soient dissemblables. L'idée principale est de définir K centroïdes.[88]

1-Fixer K le nombre de centroïde.

2-Initialiser les  $\mu_j$  Pour chaque centroïde :

- Sélectionner un pixel x au hasard parmi les pixels de la base d'apprentissage.
- $\mu_j$  est initialiser par la valeur de ce pixel.

3-Mise à jour de  $\mu_j$  façon stochastique. Pour un nombre prédéterminé d'itérations :

- Sélectionner un pixel x au hasard parmi les pixels de la base d'apprentissage.
- Trouver le centre  $\mu_{j_0}$  le plus près de ce pixel :

$$\|x - \mu_{j_0}\| = \min \{ \|x - \mu_j\| / j = 1..K \}$$

- Ajuster le centre le plus près de ce pixel :  $\mu_{j_0} = 1/2 * (\mu_{j_0} + x)$

4-Déterminer les largeurs  $\sigma_j$  tel que :

$$\sigma_j = \text{biais} * \min \{ \|\mu_j - \mu_l\| / l = 1..K, l \neq j \}$$

### **III. Implémentation:**

Nous présentons dans cette partie les différentes interfaces de notre application.

#### **III.1. Environnement matériel et logiciel de programmation :**

Notre application a été réalisée sur un PC de type Core 2 Duo 2.8 GHz, 1Go de RAM sous Windows XP.

Les langages utilisés sont: Matlab version 2008 et l'environnement Netbeans.

#### **III.2. Présentation de quelques vues :**

L'application est composée d'une fenêtre principale, elle contient des boutons permettant d'accéder aux différentes fonctionnalités.



Figure 21 : Fenêtre principal de l'application

Le Bouton « a propos » donne des informations sur le projet :



Figure 22 : Information sur le logiciel

En exécutant le bouton « entrer » depuis le menu de démarrage l'interface suivante apparaît :

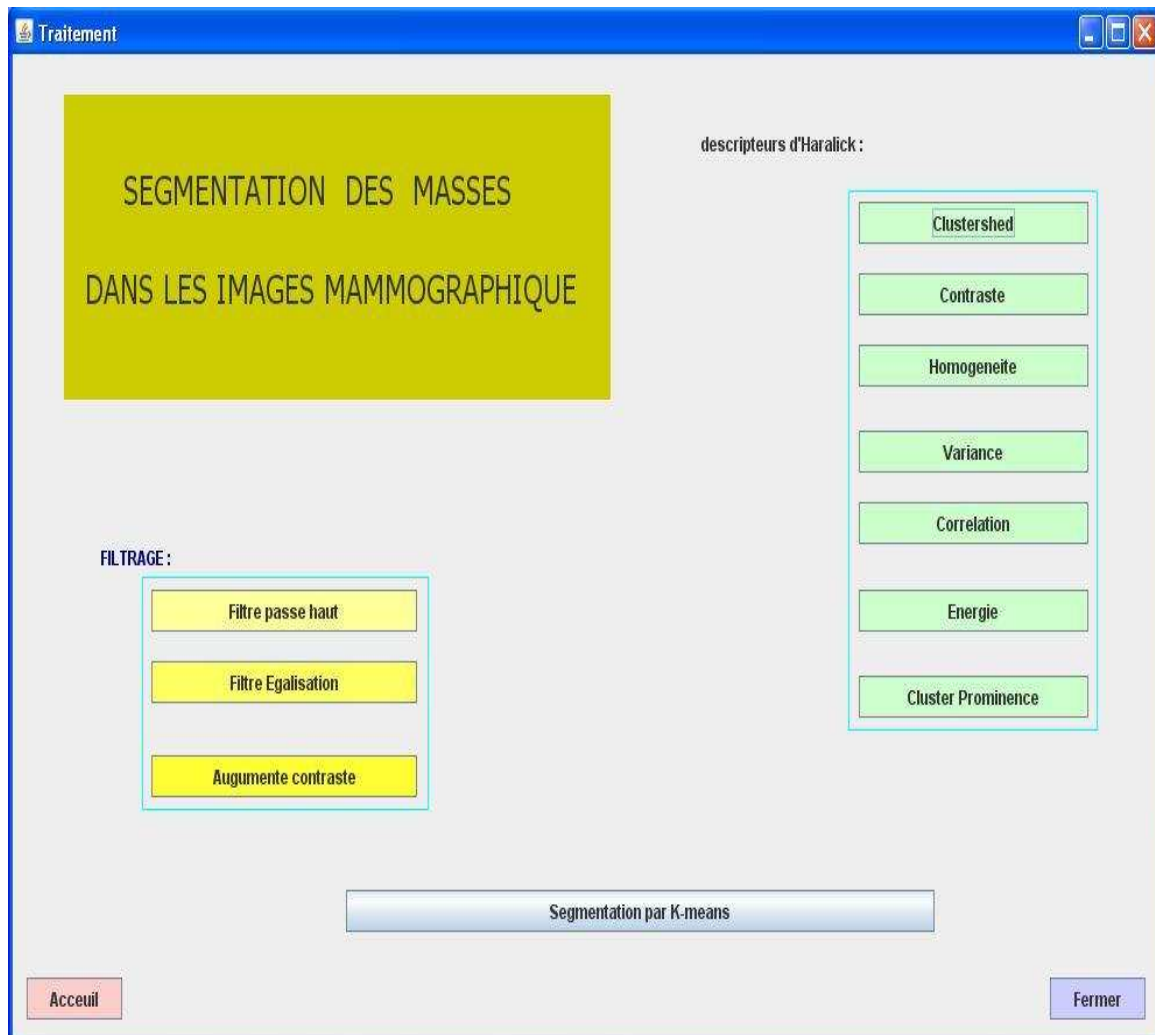


Figure 23 : fenêtre de traitement du logiciel

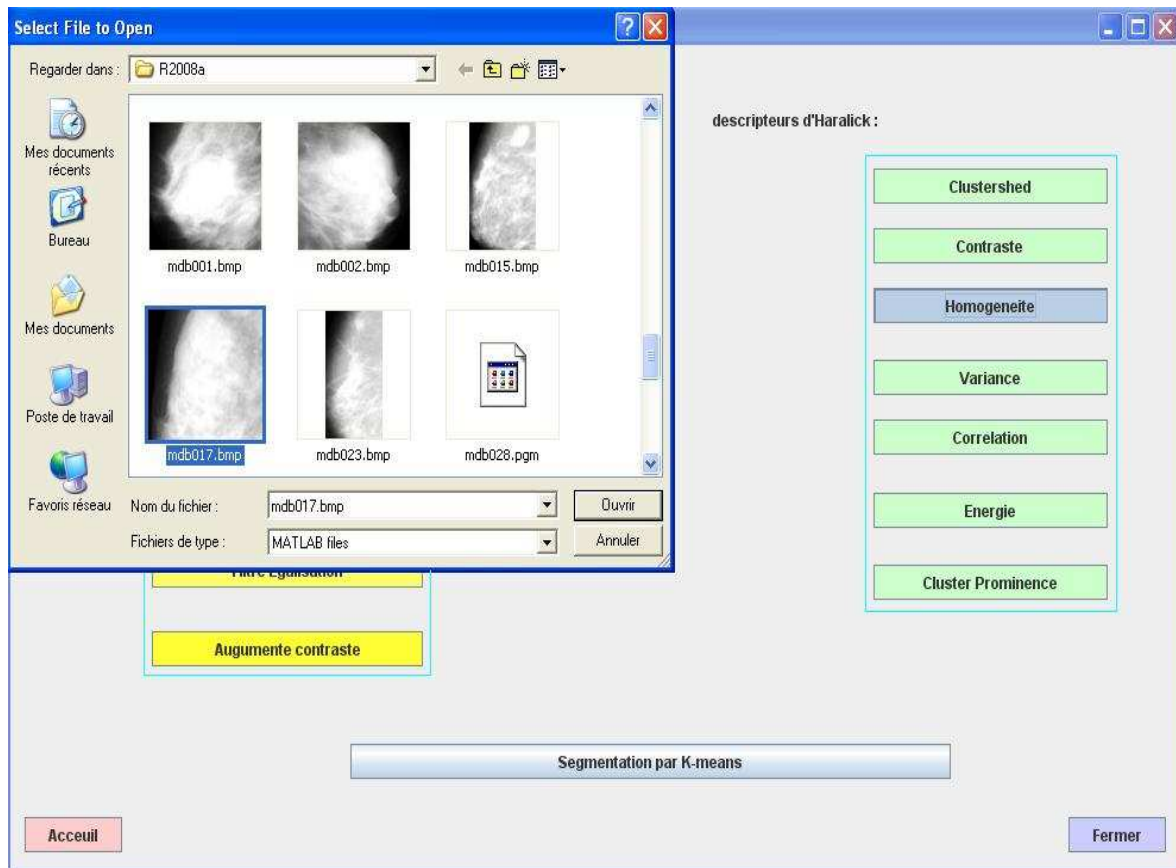


Figure 24 : fenêtre de traitement du logiciel

### Le Contraste

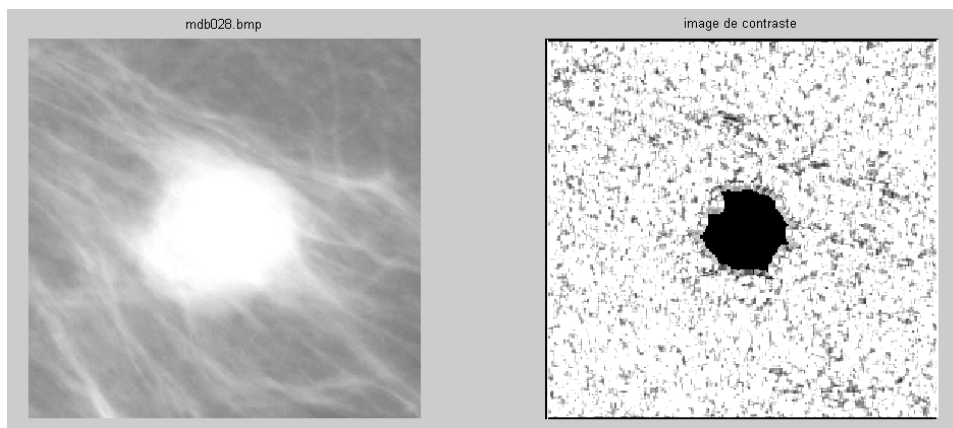


Figure 24 : fenêtre de paramètre « Contraste » de Matrice de cooccurrence (fenêtre de 3\*3).

### La Variance

- Le résultat obtenu par la variance est presque comme le résultat précédent de clusteshad, bien et clair mais les points noirs augmentent, (la couleur de l'opacité est noire et la région externe en blanc).

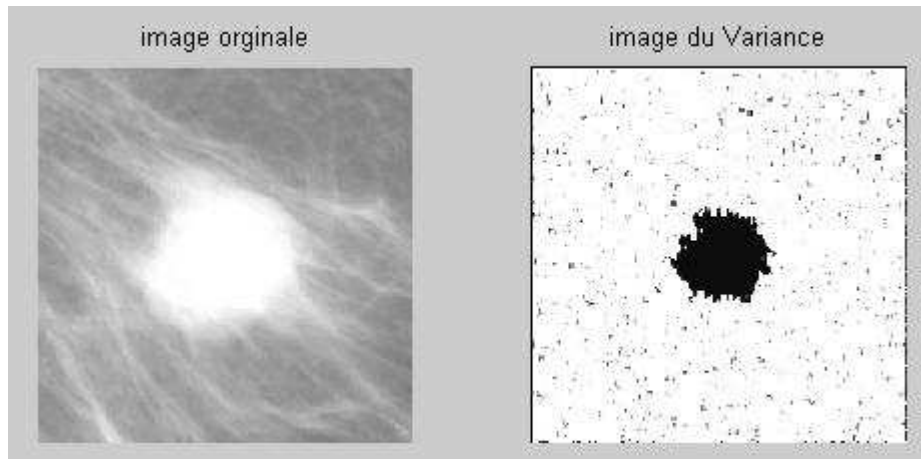


Figure 25 : fenêtre de l'image texturée, paramètre « Variance » de Matrice de cooccurrence (fenêtre de 3\*3).

### Le Clustershad

- Le résultat obtenu par le clustershad est bien et clair (la couleur de l'opacité est noire et la région externe en blanc), nous remarquons les petites points noirs qui sont des taches c.à.d. la région externe n'est pas vraiment nette.

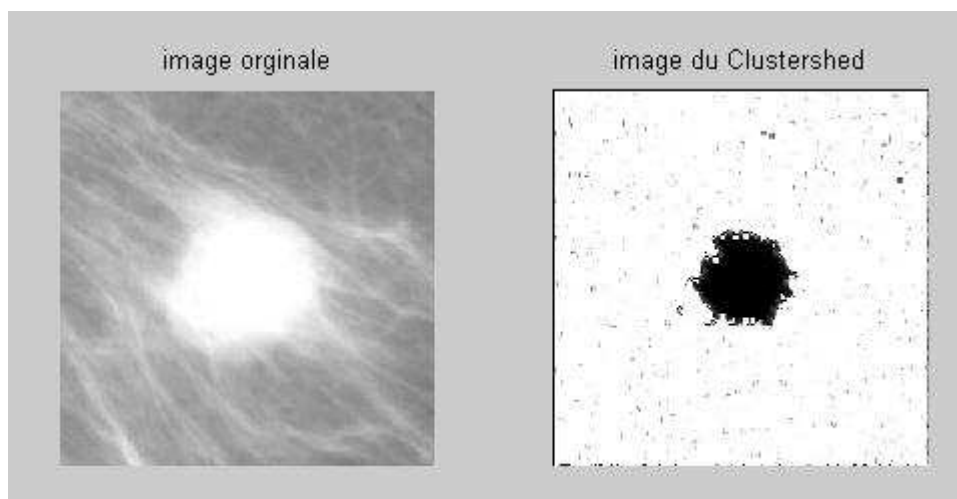


Figure 26 : fenêtre de l'image texturée, paramètre «Cluster shade » de Matrice de cooccurrence (fenêtre de 3\*3).



### ✚ La Corrélation

- Le résultat obtenu par la corrélation dans l'image mdb028 mauvaise. L'intensité est tellement faible que nous remarquons l'absence de l'opacité, et marqué la couleur noire seulement.

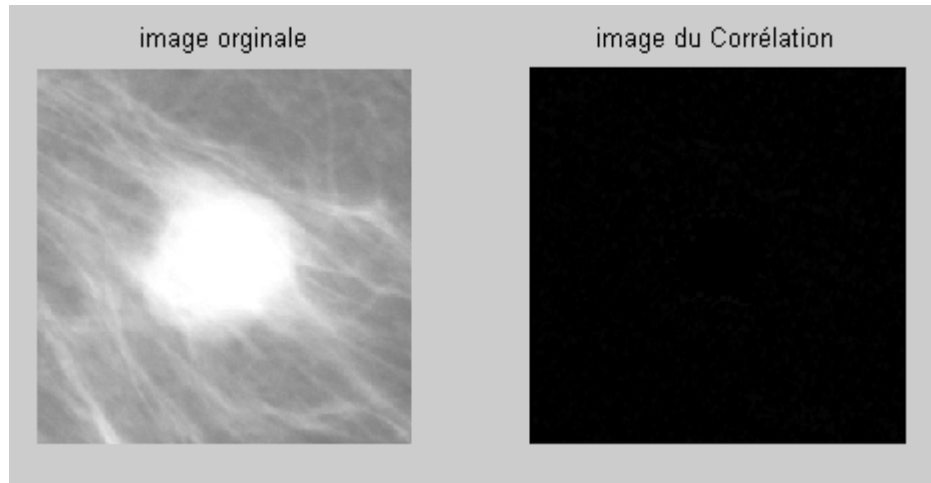


Figure 27 : fenêtre de l'image texturée, paramètre « Corrélation » de Matrice de cooccurrence (fenêtre de 3\*3).

### ✚ L'Energie

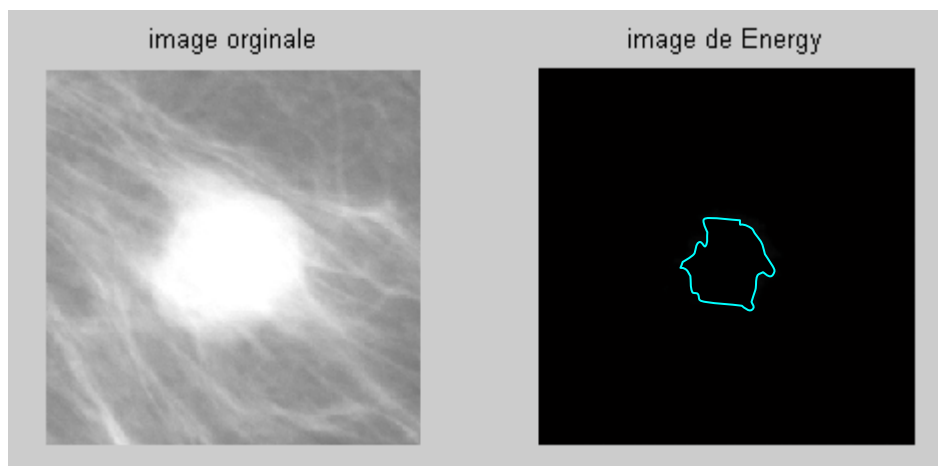


Figure 28 : fenêtre de l'image texturée, paramètre « Energie » de Matrice de cooccurrence (fenêtre de 3\*3).

### ✚ Le Cluster Prominence

- Le résultat obtenu par le cluster de prominance dans l'image mdb028 bien et très claire, l'opacité est en couleur gris et la région externe est en blanc et nette sans bruit.

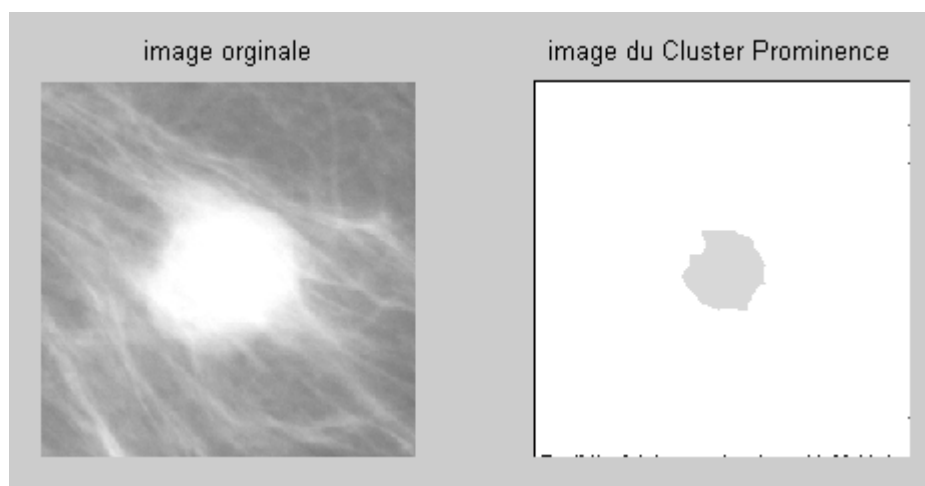


Figure 29 : fenêtre de l'image texturée, paramètre « Cluster Prominence » de Matrice de cooccurrence (fenêtre de 3\*3).

### ✚ L'Homogénéité

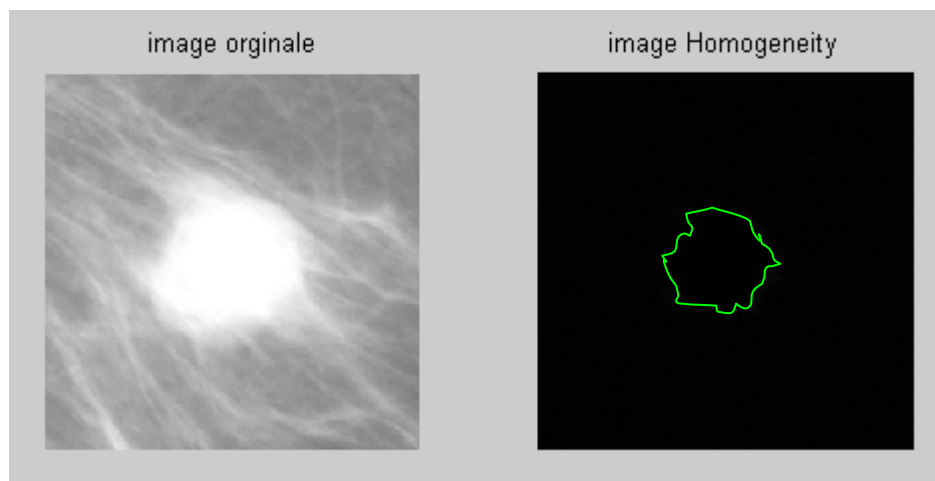
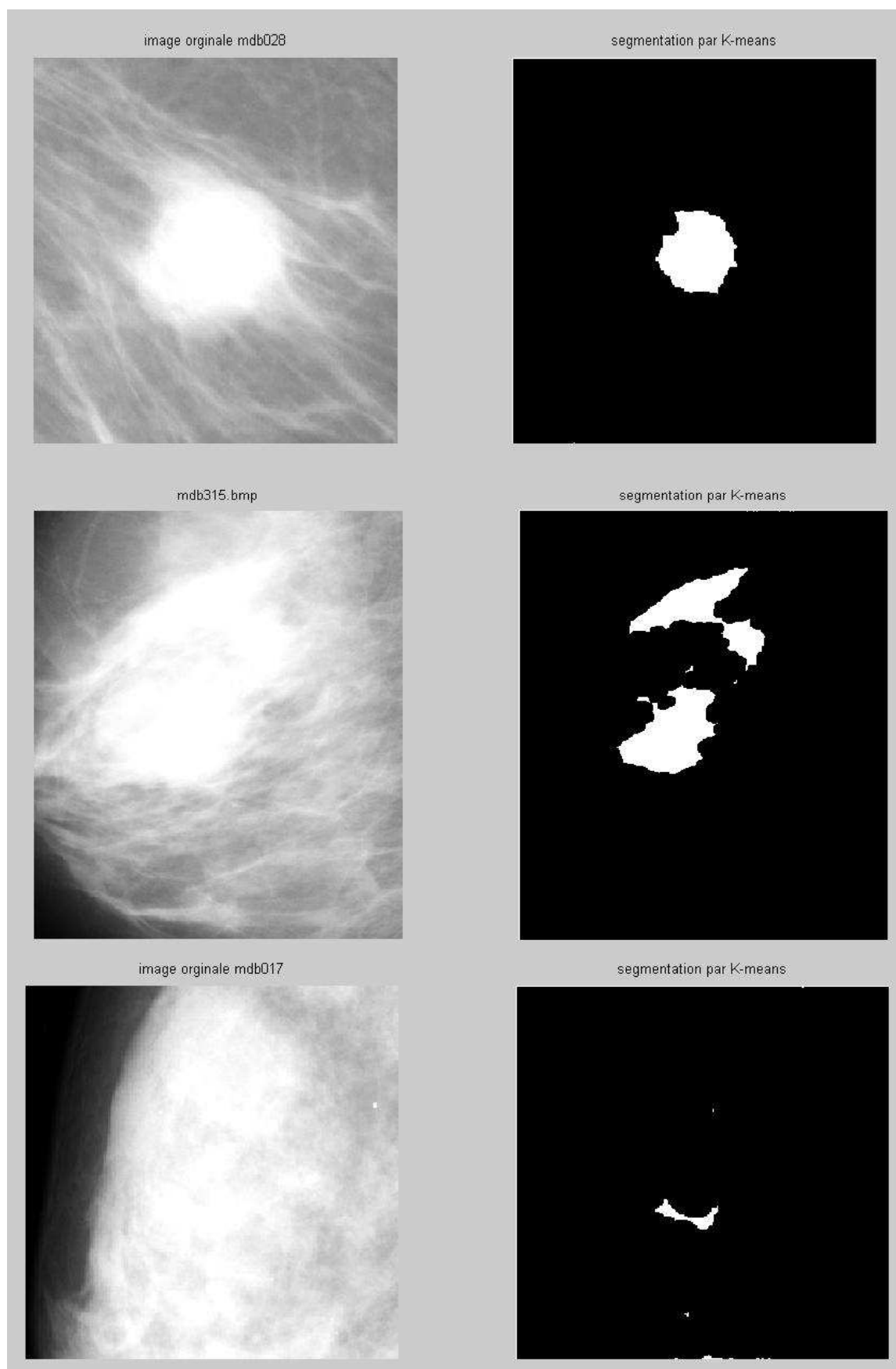


Figure 30 : fenêtre de l'image texturée, paramètre « Homogénéité » de Matrice de cooccurrence (fenêtre de 3\*3).



Figure 31 : segmentation sans calcul de descripteurs



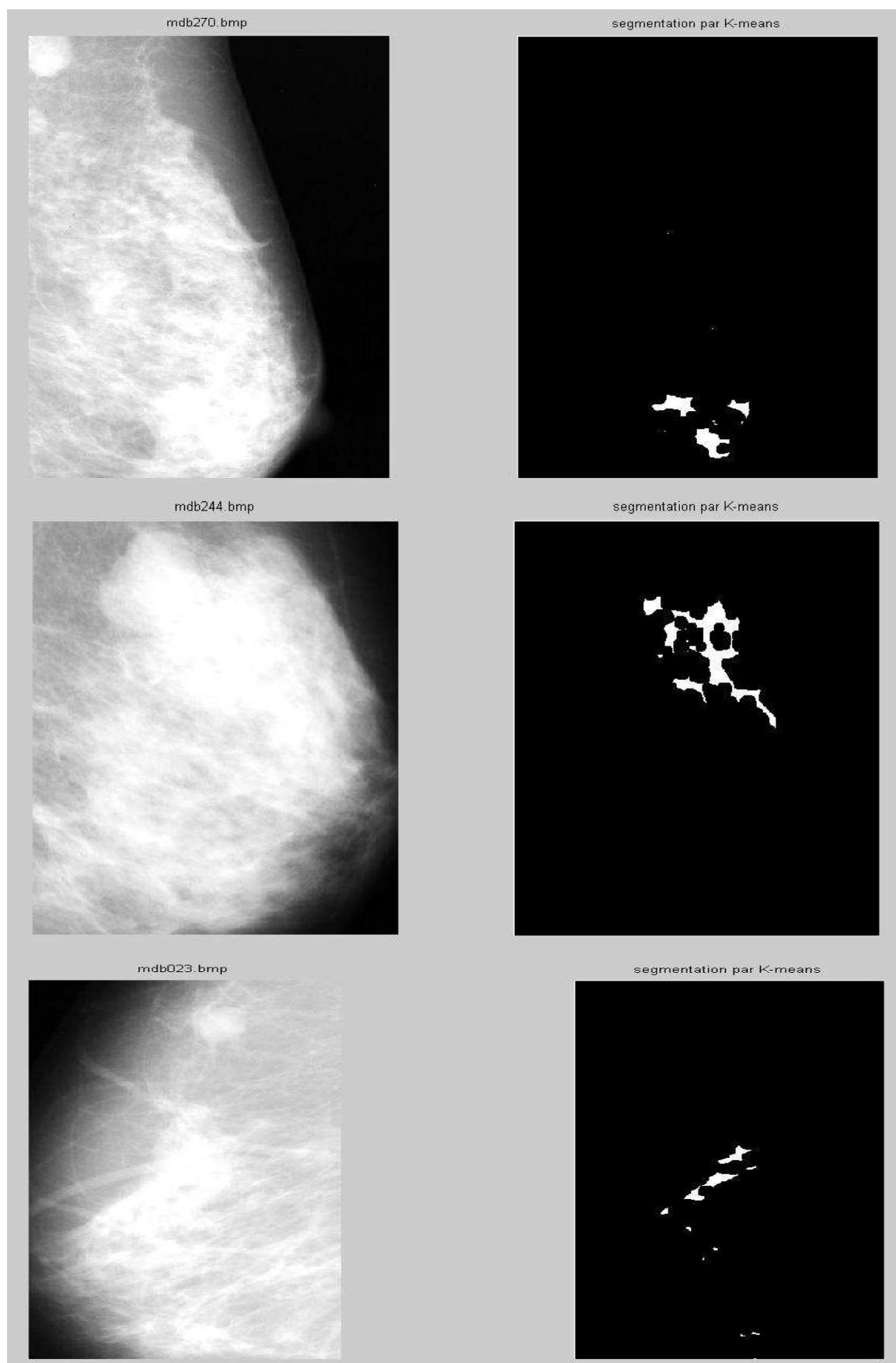


Figure 32 : L'affichage de segmentation par K-means (fenêtre de 3\*3).

### III.3. Résultats et discussions :

Après plusieurs tests nous avons choisi la taille 3x3 de la fenêtre glissante pour calculer la matrice de co-occurrence et extraire les descripteurs de texture. La matrice de co-occurrence est calculée à l'intérieur de la fenêtre, les sept descripteurs de texture est déjà indiqués dans la section II.4.

Nous avons calculés les descripteurs de texture dans la direction ( $0^\circ$ ). et pour la distance inter-pixel vaut 1.

Chaque pixel sera caractérisé par un vecteur contenant cinq descripteurs calculés, Le vecteur caractéristique d'un pixel  $i$  de l'image traitée se présente sous la forme (les descripteurs utilisés sont déterminés après plus expérimentation):

$$X_i = \{CON, ENE, CSH, CPN, CORR\} \quad (1)$$

Dans l'étape suivante, l'algorithme de clustering K-Means est appliqué en utilisant les vecteurs caractéristiques de l'image. La détermination du vecteur caractéristique est déjà présentée précédemment. Plusieurs tests ont été effectués et concernant le nombre de clusters,  $K=2$ , cluster représente le fond et l'autre la masse.

La Figure 31 présente l'application de l'algorithme K-means sans utiliser les descripteurs de texture et cela en utilisant seulement les valeurs du niveau de gris. Cette figure démontre que l'algorithme K-Means sans utiliser un descripteur n'est pas capable de faire la distinction entre les régions (l'utilisation de la valeur de niveau de gris est insuffisante).

La Figure 32 présente l'application de notre méthode qui propose un clustering K-means avec l'utilisation du descripteur de texture proposé. Cette figure démontre que l'utilisation des descripteurs de texture donnent des bons résultats, (l'image est segmentée correctement, et les clusters sont bien classés).

## IV. Conclusion

Dans ce chapitre nous avons mis au point un algorithme de segmentation très performant basé sur l'analyse de la texture par la matrice de co-occurrence et les indices d'Haralick et l'algorithme de clustéring K-means.

L'extraction des caractéristique, l'interprétation et la prise de décision sur la malignité//bénignité des masses constituent l'étape suivante du développement du module d'analyse d'image mammographique.

### Conclusion Générale et Perspectives

En analyse d'images médicales, la segmentation est aucun doute la tâche qui mobilise le plus d'efforts, elle nécessite des connaissances à priori sur le type et la qualité d'images à traité, et la structure à segmenter.

Dans ce mémoire nous avons développé un algorithme de segmentation des masses dans les images mammographiques basé sur l'analyse de la texture et l'algorithme de K-Means pour:

- Faciliter la détection des masses du sein
- L'Aider au diagnostic pour le radiologue.
- Construire un deuxième lecteur potentiel après le radiologue et réduction des erreurs commises par le medecin.

Dans le premier temps, les images mammographiques originaux ont été prétraitées afin d'éliminer le bruit et augmenter le contraste.

Par la suite nous avons prouvé la faisabilité de segmentation basé sur les indices de Haralick et l'algorithme de segmentation K-means.

Par rapport aux méthode proposées dans littérature, le grand avantage de notre algorithme est qu'il conserve la forme ainsi l'information sur le contour, qui sont des notions importantes dans le diagnostique des tumeurs mammaire ainsi que leur caractérisation.

- **Perspectives et suites possibles de ce travail...**

Ce mémoire a permis de dégager de nombreuses perspectives que nous les résumons dans ces quelques lignes :

1. Une forte liaison relie le concept segmentation et classification : après l'étape de l'extraction des paramètres pertinents une classification peut être envisagée pour d'identifier une anomalie en type bénigne ou maligne.
2. Une extension logique de cette approche est l'automatisation de système.

Bien entendu cette liste est non exhaustive et un bon nombre d'extensions supplémentaires peuvent certainement être *imaginées*...

# Annexe A

Les quatorze caractéristiques de Haralick:

- **Moyenne** =  $\frac{1}{2} \sum_{i,j} (iP(i, j) + jP(i, j))$

- **Entropy** =  $-\sum_{i,j} P(i, j) * \log(P(i, j))$

- **Correlation** =  $\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)(j - \mu_y) p(i, j)}{\sigma_x \sigma_y}$

- **Variance** =  $\sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i - MOY)^2 * p(i, j)$

- **Inertia** =  $\sum_{i,j} (i - j)^2 * P(i, j)$

- **Contrast** =  $\sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i - j)^2 * p(i, j)$

- **Energy** =  $\sum_{i=0}^{N_g} \sum_{j=0}^{N_g} p(i, j)^2$

- **Homogeneity** =  $\sum_{i=0}^{N_g} \sum_{j=0}^{N_g} \frac{p(i, j)}{1 + (i - j)^2}$



- **Cluster Shade** =  $\sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i + j - MOY)^4 * p(i, j)$
  
- **Cluster Prominence** =  $\sum_{i=0}^{N_g} \sum_{j=0}^{N_g} (i + j - MOY)^4 * p(i, j)$
  
- **Max Probability** =  $\max(P)$
  
- **Inverse Variance** =  $\sum_{i, j} \frac{P(i, j)}{(i - j)^2}$
  
- **Mode Probability** =  $\text{mod}(P)$
  
- **Moment Differential inverse** =  $\sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$

# Annexe B

## Brève description de base d'images mammographiques : « Mini-Mias »

*Dans cette annexe, nous donnons une brève description de la base d'images mammographiques qui a été utilisée le long de ce mémoire pour valider nos résultats de segmentation : la « MIAS ».*

**MIAS** : Mammography image analysis society [49] contient un ensemble de 322 images mammographiques numérisées de type de type MLO (incidence latérale oblique) qui explore le sein gauche et droite de 161 femmes.

**Les caractéristiques de la base MIAS** : les films mammographiques ont été extraits de : « *UK National Breast Screening Programme* », numérisés avec une résolution de 200 µm par pixel. La base contient aussi la localisation de différentes anomalies présentes dans une mammographie, par des radiologues experts.

Les mammographies de format (.pgm) sont nommées : « mdb xxx », dont xxx représente le numéro de l'image qui varie de 001 au 322. La taille de chaque image est 1024 Pixel x 1024 Pixel, de NG qui varie entre le 0-255.

Pour chaque mammographie, la base de données MIAS fournit des informations sur :

### 1. Le type de la densité mammaire

F Fatty

G Fatty-glandular

D Dense-glandular

### 2. Le type de la lésion si elle existe

CALC Calcification

CIRC Well-defined/circumscribed masses

SPIC Spiculated masses

MISC ill-defined masses

ARCH Architectural distortion

ASYM Asymmetry

NORM Normal

### 3. La nature de la lésion

B Benign

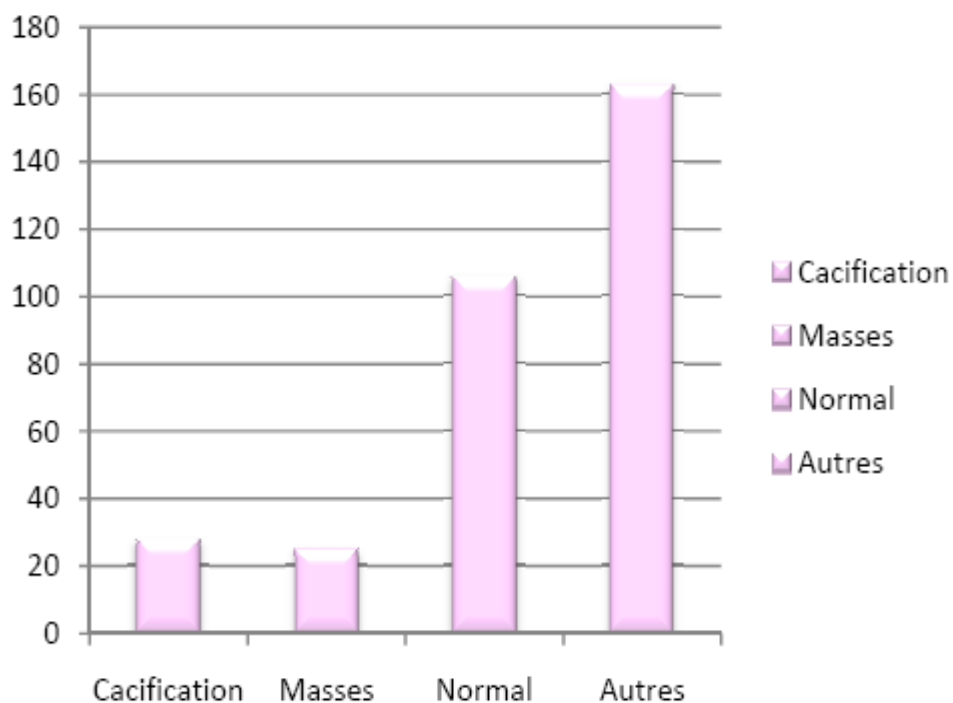
M Malignant

### 4. La localisation de lésion

-Les coordonnées de centre de la lésion(x, y), L'origine de l'image est le coin supérieur gauche.

-Rayon approximatif (en Pixel) d'un cercle enfermant l'anomalie.

Le graphe ci-dessous présente les 322 images distribuées en : 28 images qui contiennent des foyers de %cas, 25 images qui contiennent des masses, 159 images contiennent d'autres anomalies et 106 images normales (non pathologique).



**La répartition de la base de données**

## Références Bibliographiques

- [1] : Ismahen HADJII. Approche morphologique pour la Segmentation d'Images Médicales, Application à la Détection des Lésions Mammaires, mémoire pour obtenir le titre de magister en électronique biomédicale, 06 juillet 2011.
- [2] : Dr Alexandra De Grave, Pr Anne Vandembroucke, Garantir la qualité du dépistage du cancer du sein , La Revue de la Médecine Générale, n° 219, janvier 2005.
- [3] : A. travade, A. Isnard, H. Gimbergues. Pathologie mammaire. 2e édition, Paris, Masson, Imagerie Médicale, Dec 2000, 223 pages.
- [4] : C. Vachier. Extraction de caractéristiques, segmentation d'image et morphologie mathématique. These Ecole des Mines de Paris, 18 Décembre 1995.
- [5] : Andreas RICK. Représentation de la variabilité dans le traitement d'images flou, application à la mammographie numérique, thèse docteur de l'Université Paris 6, 17 décembre 1999.
- [6] R. M. Haralick. Statistical and structural approaches to texture. proceedings of the IEEE, 67:786–804, Mai 1979.
- [07]. : C. H. Chen, L. F. Pau & P. S. P. Wang, éditeurs. Handbook of pattern recognition and computer vision. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2000.
- [08] Brodatz P., *Texture : A photographic album for artists and designers*. New York : Dover 1966.
- [09] : V. Karathanassi, C. Iossifidis & D. Rokos. A texture-based classification method for classifying built areas according to their density. In International Journal of Remote Sensing, Volume 21, Number 9, pages 1807–1823. Taylor and Francis Ltd, 2000

- [10]. : Tahar Iftene & Abdelmounaine Safia. Comparaison Entre La Matrice De Cooccurrence Et La Transformation En Ondelettes Pour La Classification Texturale Des Images Hrv (Xs) De Spot. Teledetection, 2004.
- [11]: R.M. Haralick, K. Shanmugam & I. Dinstein. Textural Features for Image Classification. TransSMC, vol. 3, no. 6, pages 610-621, November 1973.
- [12] Algorithmes multidimensionnels et multispectraux en Morphologie Mathématique : Approche par méta-programmation » Thèse pour obtenir le grade de Docteur de l'école des Mines de Paris Spécialité « Morphologie Mathématique » Présentée et soutenue publiquement par Raffi ENFICIAUD le 26 février 2007
- [13] : J.P COCQUEREZ, S PHILIPP – Analyse d'images : filtrage et segmentation, Masson Paris – 1995.
- [14] S. Bres, J. Jolion, F. Lebourgeois. Traitement et analyse des images numérique. Paris, Lavoisier, Oct 2003, 411 pages.
- [15] Jean Jaques Rousselle. Les contours actifs, méthode de segmentation. Application à l'imagerie médicale. These Université François Rabelais de Tours, 09 juillet 2003.
- [16] J.M. Rendon Mancha, " Régions Actives Morphologiques : Application à la Vision par Ordinateur ", Thèse de doctorat, Université René Descartes - Paris V, (2002).
- [17] [Chia-Hung Wei](#), Computer Methods and Programs in Biomedicine [Volume 106, Issue 3](#) , Pages 234-248, June 2012
- [18] D Mbark IGGANE, Driss MAMMASS, Ali KHENCHAF and Jean-Christophe Segmentation d'images sonar par matrice de cooccurrence. International Conference: Sciences of Electronic, Technologies of Information and Telecommunications 2009 – TUNISIA.
- [19] An Adaptive K-means Clustering Algorithm for Breast Image Segmentation, Bhagwati Charan Patel, November 2010
- [20] page personnelle de Samuel Boudet : [www.samuelboudet.com/fr/matlab](http://www.samuelboudet.com/fr/matlab)
- [21] T. Bousch, Le poisson n'a pas d'arêtes, Ann. Inst. H. Poincaré Probab. Statist. 36(4), 489–508, 2000.

[22] J-P. Cocquerez, S. Phlipp. Analyse d'images : filtrage et segmentation. Masson, 1995, 457 pages.

[23] Mammography Image Analysis Society (MIAS), 2003, Disponible sur le net à l'adresse: [Http://marathon.csee.usf.edu/mammography/database.html](http://marathon.csee.usf.edu/mammography/database.html)

[24] Bong<ryul Lee, Jong<doo Lee, Myeong<jin Lee. Automated Recommendation of Initial Mass Positions for Mass Segmentation in Digital Mammograms. International Conference on Electronics and Information Engineering (ICEIE), IEEE, VOL.2, 2010.

[25] L O. Martins, G.z Junior, A.Silva+, A. C. Paiva+, M. Gattass\_, Detection of Masses in Digital Mammograms using K<means and Support Vector Machine, Electronic Letters on Computer Vision and Image Analysis 8(2):39<50, 2009

[26] S.G. Al<Kindi, Ghassan. A. Al<Kindi, Breast Sonogram and Mammogram Enhancement Using Hybrid and Repetitive Smoothing<Sharpening Technique, IEEE, 2011

## Résumé

Dans un système d'aide au diagnostic (CAD) en mammographie numérique, la détection de masse est une tâche importante et délicate pour le diagnostic précoce du cancer du sein, cette difficulté se résume, globalement dans la complexité des images mammographiques (densité mammaire) et la diversité des masses à segmenter coté contour et forme. Tenant compte de cette difficulté nous avons proposé un algorithme basé sur l'analyse statistique de la texture par les matrices de co-occurrences, les descripteurs de harralick et l'algorithme de clustering ( K-means).

**Mots Clés :** Opacité, Segmentation, Mammographie, Analyse statistique, Texture, Matrice de co-occurrence, K-means.

## Abstract

In a system for the diagnosis (CAD) in digital mammography, the mass detection is an important and difficult task for the early diagnosis of breast cancer, this difficulty is due in large part, to the complexity of mammographic images (breast density) and the diversity of the masses segmented side contour and shape. Taking account of this difficulty we have proposed an algorithm based on statistical analysis of texture by co-occurrence matrices, descriptors Harralick and K-means clustering algorithm.

**keywords :** Opacities, Segmented, mammography, Statistical Analysis, Texture, co-occurrence matrix, K-means.

### ملخص

في نظام لتشخيص (CAD) في تصوير الثدي الرقمي، والكشف الشامل هو مهمة هامة وصعبة للتشخيص المبكر لسرطان الثدي، هذه الصعوبة تتلخص في جزء كبير منه، لتعقد الصور بالأشعة (كثافة الثدي) وتنوع الكتل المجزأة إلى جانب محيط شكل والشكل. مع مراعاة هذه الصعوبة اقترحنا خوارزمية مبنية على التحليل الإحصائي للنسيج من شارك في حدوث المصفوفات، واصفات harralick، وخوارزمية المجموعات K-means

**الكلمات المفتاحية :** العتامة، مجزأة، التصوير الشعاعي للثدي، والتحليل الإحصائي، والملمس، وشارك في حدوث مصفوفة، K-الوسائل.